Neuroscience and Biobehavioral Reviews 154 (2023) 105421

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Review article

Functional imaging studies of acute administration of classic psychedelics, ketamine, and MDMA: Methodological limitations and convergent results

Sophia Linguiti [a,b,1], Jacob W. Vogel [a,b,c,1], Valerie J. Sydnor [a,b], Adam Pines [a,b,d], Nick Wellman [a,b], Allan Basbaum [e], Claudia R. Eickhoff [f,g], Simon B. Eickhoff [f,h], Robert R. Edwards [i], Bart Larsen [a,b], Andrew McKinstry-Wu [j], J. Cobb Scott [b,k], David R. Roalf [b], Vaishnavi Sharma [a,b], Eric C. Strain [l], Gregory Corder [b], Robert H. Dworkin [m], Theodore D. Satterthwaite [a,b,*]

[a] Penn Lifespan Informatics and Neuroimaging Center (PennLINC), Philadelphia, PA, United States
[b] Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, United States
[c] Department of Clinical Sciences, Malmö, SciLifeLab, Lund University, Lund, Sweden
[d] Department of Psychiatry, Stanford University, Stanford, CA, United States
[e] Department of Anatomy, University of California, San Francisco, United States
[f] Institute of Neuroscience and Medicine, (INM-1, INM-7), Research Centre Jülich, Jülich, Germany
[g] Institute of Clinical Neuroscience and Medical Psychology, Medical Faculty, Heinrich-Heine-University, Düsseldorf, Germany
[h] Institute of Systems Neuroscience, Medical Faculty, Heinrich-Heine University, Düsseldorf, Germany
[i] Department of Anesthesiology, Perioperative and Pain Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States
[j] Department of Anesthesiology and Critical Care, Neuroscience of Unconsciousness and Reanimation Research Alliance (NEURRAL), University of Pennsylvania, Philadelphia, United States
[k] VISN4 Mental Illness Research, Education, and Clinical Center at the Corporal Michael J. Crescenz VA (Veterans Affairs) Medical Center, Philadelphia, PA, United States
[l] Department of Psychiatry and Behavioral Sciences, The Johns Hopkins University School of Medicine, 5510 Nathan Shock Drive, Baltimore, MD, United States
[m] Department of Anesthesiology and Perioperative Medicine, University of Rochester School of Medicine and Dentistry, Rochester, NY, United States

ARTICLE INFO

Keywords:
Functional magnetic resonance imaging
FMRI
Psychedelics
Ketamine
Psilocybin
LSD
MDMA

ABSTRACT

Functional magnetic resonance imaging (fMRI) is increasingly used to non-invasively study the acute impact of psychedelics on the human brain. While fMRI is a promising tool for measuring brain function in response to psychedelics, it also has known methodological challenges. We conducted a systematic review of fMRI studies examining acute responses to experimentally administered psychedelics in order to identify convergent findings and characterize heterogeneity in the literature. We reviewed 91 full-text papers; these studies were notable for substantial heterogeneity in design, task, dosage, drug timing, and statistical approach. Data recycling was common, with 51 unique samples across 91 studies. Fifty-seven studies (54%) did not meet contemporary standards for Type I error correction or control of motion artifact. Psilocybin and LSD were consistently reported to moderate the connectivity architecture of the sensorimotor-association cortical axis. Studies also consistently reported that ketamine administration increased activation in the dorsomedial prefrontal cortex. Moving forward, use of best practices such as pre-registration, standardized image processing and statistical testing, and data sharing will be important in this rapidly developing field.

1. Introduction

Psychedelic drugs have been used by humans for spiritual purposes for thousands of years. Shortly following the discovery of lysergic acid diethylamide (LSD) in the 1950 s, translational studies suggested potential therapeutic uses for psychedelics (Grinspoon et al., 1979; Kyzar et al., 2017). However, in the mid 20th-century, a swift shift in public opinion and regulatory landscape limited the scientific study of these

* Corresponding author at: Penn Lifespan Informatics and Neuroimaging Center (PennLINC), Philadelphia, PA, United States.
E-mail address: sattertt@pennmedicine.upenn.edu (T.D. Satterthwaite).
[1] Contributed equally as first authors

https://doi.org/10.1016/j.neubiorev.2023.105421
Received 4 May 2023; Received in revised form 13 September 2023; Accepted 2 October 2023
Available online 5 October 2023

psychotropics for several decades (Kyzar et al., 2017). Recently, a new wave of interest in psychedelics – in particular for treatment of psychiatric disorders including major depression (Pearson et al., 2022; Carhart-Harris et al., 2021) and PTSD (Mitchell et al., 2021) – has emerged in tandem with regulatory reform (Siegel et al., 2023), prompting an upsurge in clinical trials and translational research.

As part of efforts to understand mechanisms of psychedelic action in humans, there has been a proliferation of studies examining psychedelics with non-invasive functional imaging methods. Typically, this is done using functional MRI (fMRI), which measures real-time dynamics of neural activity linked to blood oxygen in the brain. fMRI offers a valuable tool for discovering associations between psychedelic administration and changes in brain activity; these can inform mechanistic theories of psychedelic action in the human brain. Prevailing theories posit that psychedelics enhance plasticity (Siegel and Nicol, 2023; Ly et al., 2018), change activity in brain regions with high expression of monoamine neurotransmitters (Vollenweider and Smallridge, 2022; Tagliazucchi et al., 2016), and facilitate integrative processing across large-scale brain networks (Girn et al., 2023). Due to its relative ease of acquisition and non-invasive nature, fMRI is well-suited to help scientists understand the impact of psychedelics on human brain function in vivo. Specifically, the whole-brain coverage and good spatial resolution of fMRI faciliates localization of psychedelic effects to specific brain regions or networks. As such, there is the hope that fMRI could over time yield a quantitative biosignature of circuit engagement that could be used in both clinical trials and to facilitate the development of novel therapeutic agents (Wager et al., 2013; Carmichael et al., 2018; Sadraee et al., 2021).

As a result, studies of acute responses to psychedelics increasingly have incorporated fMRI. However, as for many other emerging technologies, translational imaging research using fMRI has historically witnessed waves of excitement regarding the promise of the technology, followed by sober retrenchment as methodological challenges have been increasingly appreciated. Two of the most painful methodological challenges fMRI research has faced are the substantial risk of false positive results (Eklund et al., 2016; Marek et al., 2022; Botvinik-Nezer et al., 2020; Bennett et al., 2009) and the confounding influence of motion artifact (Power et al., 2012; Satterthwaite et al., 2012, 2019). First, false positive results (Type I error) are a major threat to the interpretations of fMRI studies. These studies often include whole-brain analyses that involve statistical tests across hundreds of thousands of voxels. Methodological standards for appropriate control of the many statistical tests typically used in fMRI research have evolved in response to compelling research demonstrating that once-popular methods are, in fact, vulnerable to generating false-positive findings (Poldrack et al., 2008, 2017; Woo et al., 2014; Roiser et al., 2016). Alongside such shifts in standards, sample sizes of typical fMRI studies have expanded dramatically, in response to the greater statistical power required after appropriate correction for multiple testing.

Another notorious methodological consideration in fMRI research is motion artifact. Due to spin-history and partial volume effects, in-scanner motion can have a dramatic effect on the measured fMRI signal at a given location, as well as on downstream derived measures. Several studies have documented the impact of motion in both task (Friston et al., 1996; Siegel et al., 2014) and resting-state fMRI (Power et al., 2012; Satterthwaite et al., 2012; Van Dijk et al., 2012), a major source of bias in many different subfields, including development, aging, and translational studies in clinical populations (Makowski et al., 2019; Turesky et al., 2021). However, thus far, the impact of motion on fMRI studies of psychedelics has received little attention, despite the clear possibility that acute administration of psychedelic agents might lead to systematic differences in participant motion during scanning. Importantly, such methodological concerns may interact: the large effect size of in-scanner motion artifact on some measures has the potential to drive results, especially when combined with inadequate methods for Type I error control.

In this context, we conducted a review of fMRI studies that examined the acute administration of psychedelics in humans. Our original goal was a meta-analysis, but we found this to be infeasible given the heterogeneity of the current literature. As such, we here present the results of a systematic review. We considered both task-based studies designed to study activation as well as studies that examined intrinsic functional connectivity. Specifically, we searched the literature for fMRI studies that monitored acute responses to classic psychedelics (agonists of the 5-HT2A receptor) including LSD and psilocybin. However, to ensure broad coverage of the field, we also considered atypical psychedelics that have been the subject of clinical interest in psychiatry, including ketamine and methylenedioxymethamphetamine (MDMA). Cannabinoids were not included due to substantially differing mechanisms of action and inconsistent psychedelic properties. As described below, we first summarized the study design, sample size, and sample uniqueness (in the context of other published reports). Next, we evaluated how each study addressed motion artifact and corrected for Type I error, comparing the reported methods to contemporary best practices. Finally, having identified a sub-sample of studies that were at least generally aligned with current methodological standards, we sought to synthesize the reported findings.

2. Methods

2.1. Methods overview

This study began as a pre-registered systematic review and coordinate-based meta-analysis that sought to examine the impact of acute administration of psychedelic agents on brain activation. The review was planned using the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting guideline (Moher et al., 2015) and pre-registered using PROSPERO on May 5th, 2020 (https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=183416). However, we found that the available literature was too heterogeneous to support a meta-analysis. Furthermore, on full-text review, we observed that many of the studies were small, often appeared to use the same samples, did not appear to account for the potential impact of motion artifact, and did not adhere to contemporary field-standard guidelines for Type I error control. Accordingly, we re-framed our initial search to systematically examine design characteristics, sample size, sample uniqueness, handling of in-scanner motion, and Type I error control of all fMRI studies that investigated the acute administration of psychedelics – including both task fMRI and intrinsic connectivity. This approach allowed us to identify a sub-sample of studies that appeared to use contemporary methods for control of motion artifact and Type I error. These studies are discussed as part of a scoping, integrative review. Documentation of methodological outline of the systematic review can be seen in Fig. S1.

2.2. Original search criteria (for meta-analysis)

The following was our initial criteria for the meta-analysis, which was later revised for a systematic review (see below). A literature search of fMRI studies that monitored the acute administration of psychedelics published between January 1, 1990, to May 8, 2020, was conducted using pre-registered search terminology on PubMed/MEDLINE, EMBASE, Web of Science, Cochrane Library, and PsycINFO. We used the the following search terms: (“MRI” OR “magnetic resonance imaging” OR “fMRI” OR “BOLD” OR “brain mapping”) AND (“psychedelic” OR “psychoactive” OR “hallucinogen” OR “ayahuasca” OR “DMT” OR “N,N-dimethyltryptamine” OR “5-MeO-DMT” OR “5-methoxy-N,N-dimethyltryptamine” OR “LSD” OR “lysergic acid diethylamide” OR “psilocybin” OR “4-hydroxy-N,N-dimethyltryptamine” OR “psilocin” OR “mescaline” OR “peyote” OR “2-(2-Chlorophenyl)− 2-(methylamino)cyclohexanone” OR “ketamine” OR “esketamine” OR “MDMA” OR “3,4-Methylenedioxymethamphetamine” OR “ecstasy” OR “ibogaine” OR

"scopolamine" OR "2,5-dimethoxy-4-iodoamphetamine" OR "DOM" OR "2,5-dimethoxy-4-methylamphetamine" OR "DOB" OR "2,5-dimethoxy-4-bromoamphetamine" OR "4-bromo-2,5-dimethoxyphenethylamine" OR "2 C-B" OR "2,5-dimethoxy-4-iodophenethylamine" OR "2 C-I" OR "PCP" OR "phencyclidine" OR "salvia" OR "salvia divinorum" OR "DXM" OR "dextromethorphan"). Additionally, the search criteria were limited to publications written in English.

All results from the above query were organized using Covidence; details were extracted and duplicate references were removed. In total, 3666 abstracts were screened for the initially proposed meta-analysis. To evaluate meta-analytic study inclusion we used the following criteria: abstracts were included if they (1) examined the acute administration of psychedelics to participants completing fMRI and (2) the analysis included a "psychedelic administration vs. baseline" contrast. Abstracts were excluded based on the following criteria: (1) sample size $n < 10$ subjects; (2) primary imaging modality was not BOLD fMRI (e.g., PET); (3) did not use an experimentally-induced administration of psychedelics (e.g., examined long-term or chronic use); (4) primary drug was a cannabinoid receptor ligand (e.g., THC from cannabis); (5) field of view and reported analyses did not cover the whole brain; (6) did not report results in standard stereotaxic reference space coordinate system (MNI or Talairach); (7) did not report statistics in sufficient detail to allow for robust cluster identification (e.g., voxel height of $p < 0.001$ (uncorrected) and/or cluster-corrected $p < 0.05$). Lastly, to avoid inaccurately limiting our included studies, all publications that did not clearly report whether they met exclusion criteria were included in the full-text screening. In total, 502 full-text publications were reviewed by two independent reviewers. Selection criteria were further refined based on issues that emerged during the review process. Specifically: (1) papers reporting $>$ 72 h between drug administration and fMRI were excluded; (2) papers with participants aged $<$ 18 years old were excluded.

Following this review, it became clear that the included studies were too heterogeneous to conduct a quantitative meta-analysis. Only five papers clearly met all inclusion criteria; 390 were excluded, and 107 were marked as ambiguous. This ambiguity largely came from three obstacles: (1) many studies were resting-state (i.e., task-free) functional connectivity studies without an explicit drug vs. placebo activation contrast, (2) publications did not clearly report whether whole-brain analyses were completed, and (3) the duration of "acute administration" differed widely among studies. In addition to design heterogeneity, we also noted that there was substantial methodological heterogeneity in the existing literature. Based on a simulation study, 17–20 experiments are needed to complete an ALE meta-analysis with sufficient power (Eickhoff et al., 2016). The five publications yielded by our search was therefore deemed insufficient for a quantitative meta-analysis. Accordingly, we re-framed the goal of this study from a meta-analysis into a systematic review. Three stages were defined. First, we refined our search once again to be more permissive, allowing for a wide survey of psychedelics fMRI literature. Second, we evaluated the existing literature with regard to variation in key design and methodological parameters outlined below. Finally, we conducted a qualitative, integrative review of the sub-set of studies that used methods that align with minimum contemporary field-standard practices.

2.3. Revised search criteria (for systematic review)

After the goals of the study were re-framed, we adjusted the inclusion criteria of our search to allow for (1) both patient and control participant studies, (2) both resting and task-based fMRI, (3) relaxing of earlier criteria regarding statistical correction, whole-brain coverage, and reporting of MNI coordinates. From the original 502 papers included in the full-text review, 98 publications (59 ketamine, 17 lysergic acid diethylamide (LSD), 14 psilocybin, and eight 3,4-Methylenedioxymethamphetamine (MDMA)) were judged to meet the updated inclusion criteria and were included in the systematic review. We also found four studies of ayahuasca and two studies of N,N-dimethyltryptamine (DMT). Three of the four ayahuasca studies included participants that used ayahuasca regularly over several years for religious purposes, which we judged as a mixed acute/chronic study design that was distinct from other studies included in this review. We therefore excluded these three studies. The three studies focusing on DMT/ayahuasca were also not considered given the small number remaining.

2.4. Assessment of study design

For each study reviewed, we recorded the psychotropic used, dosage, method of drug administration, whether the sample was composed of healthy individuals or patients, and the interval between drug administration and scan. For the latter, we converted all scores to minutes from scan, such that negative numbers indicated minutes before scanning began and positive numbers indicated minutes after scanning began. For the study population, when a clinical sample was studied, we recorded participant diagnostic status (e.g. major depressive disorder, schizophrenia, etc.). We grouped studies employing explicit tasks into broad task-based categories, including attention, emotion, language, learning/memory, music, pain, rest, reward, social, visual, working memory. We also categorized the type of analysis performed, including whether studies utilized activation vs. connectivity analyses. We evaluated characteristics of the analytic strategy, including whether analyses were whole-brain or focused on specific ROIs or networks; we also recorded the seed used in seed-based connectivity analyses of functional connectivity. Finally, as described in detail below, we evaluated whether each study appeared to be from a unique sample of participants, what procedures for Type I error control the study used, and how the study attempted to control for in-scanner motion.

Notably, all study characteristics evaluated were assessed by two independent reviewers (SL & JV), and if there was disagreement, a third senior reviewer (TS) also assessed the study. Consistent with our review's emphasis on reproducible reporting procedures, we only assessed published information, and did not contact study authors if reported methodology was unclear. Whether or not studies performed appropriate correction for motion or multiple comparisons, failing to report this information unambiguously does not align with current best practices. We emphasize that the evaluation that resulted from our review and is described below should *not* be considered a condemnation of the methodological rigor of specific studies. Standards have continued to evolve over time and many studies describe their approach with contemporaneous citations. We sought to evaluate a set of factors that have been shown to limit the reliability or reproducibility of fMRI results. Additionally, we would like to emphasize that we consider the criteria outlined below to be minimal; in many respects, the evaluation standards used were liberal rather than overly stringent.

2.5. Evaluation of sample uniqueness

Beyond basic study design characteristics, we also evaluated the uniqueness of each study sample. The publication of multiple reports from the same (often small) set of participants has the potential to reduce the generalizability of results, which can be particularly problematic if not reported transparently. Sample uniqueness was assessed across studies by directly comparing the reported drug dosage, drug administration route, clinical diagnoses, recruitment strategy, study authors, the demographic breakdown of the study sample (e.g. age and sex), and the number of participants included before and after quality assurance. While some studies explicitly reported that the same sample was used, this was often not the case. For primary analysis, we constructed a binary score for uniqueness, where a sample was not considered unique if there was evidence that a previous publication had reported the study sample. However, a scale from 1 to 3 was used to indicate confidence in our assessment of participant overlap and is reported in the supplementary information (Table S1). A score of 1 (very

confident) was given if a study explicitly stated that the sample overlapped with another study. A score of 2 (high confidence) was given if this was not explicitly reported, but other characteristics (listed above) provided fairly unambiguous evidence for sample overlap. A score of 3 (moderate confidence) was given if some aspects of the study sample likely overlapped with another study, but some degree of ambiguity was present. As an example, we often encountered samples with nearly the exact same characteristics but with slightly different sample sizes (perhaps due to task-specific exclusion), and these were considered to be using the same sample.

2.6. Assessment of Type I error control

As noted above, inadequate Type I error control has been repeatedly documented as a major confound in imaging research. We evaluated procedures for multiple comparison correction according to a relatively permissive interpretation of contemporary recommendations. Each paper was given a binary score ("1"= adequate). Specifically, for whole-brain mass univariate analyses, we examined if a study used a height threshold of at least $p < 0.001$ uncorrected *and* a corrected cluster significance of at least $p < 0.05$ (Eklund et al., 2016; Poldrack et al., 2008; Woo et al., 2014; Roiser et al., 2016; Carter et al., 2016). Notably, we did not evaluate the approach or specific implementation used for the calculation of cluster significance (Eklund et al., 2016). Given prior reports of heterogeneity among methods, this approach is likely to be liberal rather than conservative. If a study reported a cluster correction but did not clearly report the approach for both cluster height and significance, this was considered inadequate. The only exception to this evaluation was when approaches that build height thresholding into cluster correction were used, including the threshold-free cluster enhancement (Smith and Nichols, 2009) or network-based statistic (Zalesky et al., 2010). Furthermore, in order to attempt to accommodate shifts towards more stringent recommendations over time, we considered Type I error correction adequate if a study performed inadequate height thresholding (e.g. $p < 0.01$) but the study also reported a table with z-values and any of the "significant" clusters exceeded $z = 3.1$ (e.g. it would have survived height thresholding of [$p < 0.001$]). If a study reported results using a combination of both an appropriate and a more liberal statistical thresholding, the study was still awarded a 1. If a study did not report a height threshold, but tables or figures (e.g. color-bar thresholding) unambiguously suggested a critical value threshold of 0.001, the study was given a 1. Finally, it should be noted that positive results were not required: a study that reported no effects surviving adequate statistical correction would still receive a score of 1.

Importantly, the above criteria only applied to mass-univariate analyses across voxels or vertices. Studies that only evaluated voxelwise testing within specific "a priori" regions of interest (with or without small volume correction) were assigned a score of 0, unless there was clear evidence of pre-registration. However, studies were not given a 0 if mass-univariate analyses were restricted to a mask derived from activation or connectivity. For seed-based analyses, although best practices would generally dictate a correction for the number of seeds tested, we did not assign a 0 to studies that did not perform this correction.

For studies performing mass univariate testing without an explicit voxel- or vertex-wise spatial component (e.g. edge-wise connections in a network, ROIs in an atlas, etc), studies using FDR, Bonferroni or equivalent correction received a score of 1. Finally, studies evaluating global activation or connectivity with a single (e.g. omnibus) measure or model were given a score of 1. Note that studies were scored for statistical correction only on contrasts evaluating the main effect of a psychedelic agent vs placebo or other control. Even if other analyses that did not meet the criteria here were reported, as long as a whole-brain contrast was evaluated and met the criteria described above, Type I error was considered adequate. As such, this evaluation yielded a relatively liberal assessment of Type I error control.

2.7. Assessment of in-scanner motion

Although motion correction is multifaceted, we used a simple two-pronged assessment. First, we evaluated whether the study disclosed criteria for the exclusion of study participants due to head motion or artifacts related to head motion. If the study reported exclusion of participants for motion, and/or reported criteria and explicit evaluation for such exclusion, the study received a score of 1 (whether or not any participants were actually excluded). Otherwise, the study received a score of 0 on this criterion.

Second, we evaluated whether the study reported any indication of addressing motion as part of statistical hypothesis testing (e.g. at the group level). Several studies have shown that strong motion effects can still be detected at the individual or group level even after preprocessing (Satterthwaite et al., 2019; Ciric et al., 2017; Siegel et al., 2017). Given known strong relationships between motion artifact and measures of both activation and connectivity, when motion systematically covaries with the independent variable, it is important to control for in-scanner motion during statistical testing. Therefore, we awarded a study a score of 1 if the study either a) included a motion summary measure (e.g. frame displacement, derivative of root mean squared variances [DVARS], etc.; number of excluded frames alone was not sufficient) as a nuisance covariate in second-level analysis, OR b) reported an evaluation of motion differences across treatment groups, OR c) reported an analysis evaluating whether a significant effect of motion was present in regions/networks/edges showing treatment effects. If none of the above scenarios were reported, the study was given a score of 0 on this criterion. The lone exception involved what we refer to as "within-run infusion" studies, where the drug of interest was infused in the middle of fMRI scanning, and modeling was only performed at the first level of inference (i.e. within-subject one-sample t-test). For all studies, these two evaluation criteria (i.e., inclusion criteria and hypothesis testing) were summed to create an overall motion score.

2.8. Summary of methodological review

For all three scoring procedures (two motion, one statistical inference), studies were assumed to have scored a 1 first and were only scored 0 if evidence supporting a score of 1 could not be found. In order to integrate across our assessment of methods for control of motion artifact and Type I error, we aggregated them into a combined score. We use this aggregated score to investigate the change in these methodological practices over time. We also acknowledge that motion artifact is often considered less of an issue in task-based activation studies. Therefore, we stratify assessments of motion and Type I error correction by paradigm (i.e. task vs. rest).

For the qualitative review portion of the manuscript, we decided to focus on studies that both demonstrated adequate control of Type I error and reported any strategy of mitigating motion artifacts. This is by no means an exhaustive criteria for methodological rigor; however, studies that did not meet at least these criteria have the potential to skew results. The results of this qualitative analysis of the selected publications appear in the Results section of this paper and are further contextualized in the Discussion.

3. Results

3.1. fMRI studies of psychedelics are generally small and highly heterogeneous

We systematically reviewed 98 fMRI studies examining the effects of the acute administration of psychedelics on BOLD activity or connectivity. The studies reviewed tended to use small sample sizes, and we found a high degree of heterogeneity in design and analytic approach. The majority of studies (91.9%) used a within-subject (cross-over) design, and the mean N for these studies across all drugs was 18.9

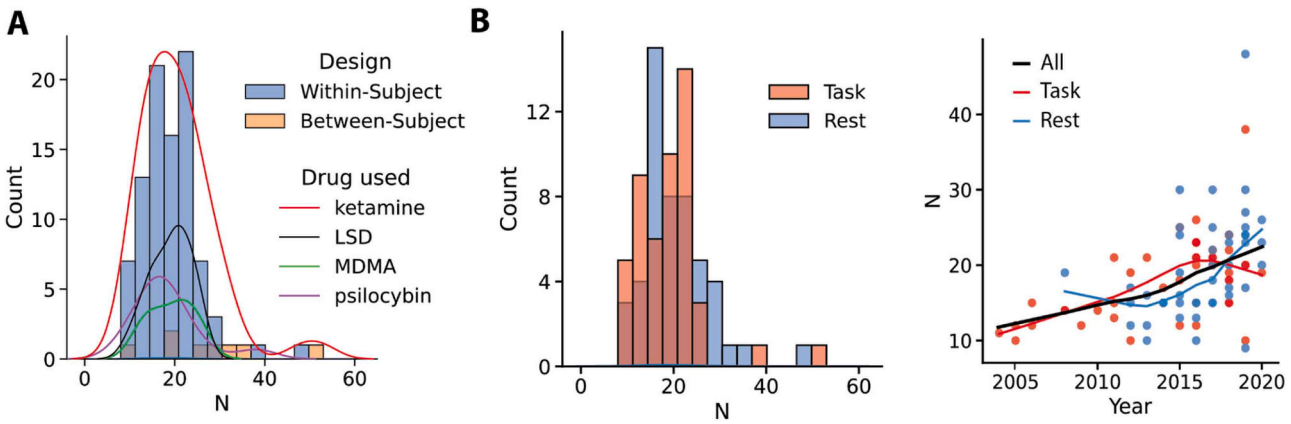

Fig. 1. Sample size across studies using fMRI to study the acute effects of psychedelic drugs. A) Histograms show the distribution of sample size (N) across all studies reviewed. Histograms are colored according to whether studies used a within-subject (blue) or between-subject (orange) design in assessing the main effect of (psychedelic) treatment. Overlaid kernel density estimate plots represent the distribution of sample size across studies examining the effects of specific drugs on the BOLD signal. B) Overlaid histograms indicating the distribution of N across resting-stage (blue) and task-based (red) studies, respectively. C) Sample size increased over time between 2004 and 2020. This plot only includes studies with within-subject design (92% of studies reviewed), and only includes studies meeting inclusion criteria (e.g. N < 10 studies excluded). The black line indicates the trend line across all studies, while the blue and red lines indicate the trend for resting-state and task-based studies, respectively.

(SD=6.3; Fig. 1**A**). This average sample size did not vary significantly across drugs, ranging from 18.0 (SD=7.4) for psilocybin studies to 19.1 (SD=6.9) for ketamine studies (Fig. 1**A**). Between-subjects designs were only seen among ketamine studies, where the average total between-subjects sample size across eight studies was 27.5 (SD=13.3). The average sample size of task-based studies ranged from 17.5 (SD=4.2) for MDMA to 21.4 (SD=8.6) for psilocybin, while the sample size of resting-state studies ranged from 14.0 (SD=8.6) for psilocybin to 21.7 (SD=5.8) for MDMA (Fig. 1**B**). There was no significant main effect of drug or task design (task or rest) on sample size, nor was there an interaction. Between 2004 and 2020, sample size increased over time (Fig. 1**C**). Note that the sample sizes reported here are calculated after removing 16 studies with N < 10 subjects as part of the inclusion criteria for our review, resulting in a clear inflation of mean sample size.

Beyond the small sample size of most studies, we observed marked heterogeneity in study design (Fig. 2, Table S1). For MDMA and ketamine, the method of administration was consistent with that typically used for clinical trials and treatment purposes. However, among psilocybin and LSD studies, method of administration varied among publications: seven LSD studies involved intravenous administration while 10 involved oral administration. Psilocybin studies were similarly split with seven involving intravenous administration and seven involving oral dosing. Published studies examined both psychiatric patient and healthy control populations; 84 papers studied healthy controls and 14 papers studied patients. Among studies that included patients with a diagnosed psychiatric disorder, 13 focused on individuals diagnosed with major depressive disorder (MDD) and one focused on individuals with schizophrenia; two of these studies used mixed samples of both patients and controls. Studies also varied in dosage and the duration between drug administration and imaging procedures (Table S1). Notably, this did not reflect use of different psychedelics. For example, ketamine administration ranged from 0.1 to 0.5 mg/kg, and studies varied in whether this amount was constant or whether a bolus was injected first, followed by a smaller amount over time. This is notable given these dosages range from subclinical to clinical. Meanwhile, timing ranged from 24 h before scanning to seven minutes after scanning onset, with the median timing being 25 min before scanning onset. These parameters are relevant, as several papers reported different results in the same study based on both dosage (e.g (Chen et al., 2019)) and when scanning occurred (e.g (Javitt et al., 2018; Li et al., 2020; Preller et al., 2020)).

Just over half of the studies reviewed (n = 50) acquired fMRI data at rest; the remaining used a wide variety of fMRI task paradigms that spanned 11 broad categories (n = 43), or used both task and rest paradigms (n = 5). In terms of analytic approach, 44 publications used activation-based analyses, 43 used connectivity, six reported results from both connectivity and activation analyses and five publications evaluated other derived measures. Finally, for studies that used seed-based analyses, there was large variety in the seed region chosen; some seed regions within the ketamine literature (e.g., dorsolateral prefrontal cortex and posterior cingulate), however, were more common.

3.2. Sample overlap in published studies is common

A theme that clearly emerged when reviewing the 98 publications included in the systematic review was that the majority published on study samples that had been included in at least one other fMRI study of psychedelics (Fig. 2, Table 1). Of the 98 publications, there were only 51 unique samples. Ketamine showed the most unique samples with 59 total publications and 37 unique samples. However, the other psychedelics examined showed more non-unique samples (Table 1). Specifically, LSD studies included 17 papers with only four unique samples, MDMA studies included eight studies with four unique samples, and psilocybin studies featured six unique samples over 14 total studies. It should be emphasized that, while many studies analyzed data from the same participants, the publications themselves often focused on separate tasks and different scans within the study protocol. Notably, among non-unique samples, 25% of studies (19 total) did not explicitly report that they used a previously published sample.

As a given study sample formed the basis of many reports, the number of unique subjects across psychedelics studies is much smaller than the total number of subjects reported across studies. All studies included in this review combine to describe data across a total of 2011 subjects if summed naively. However, after examining overlapping cohorts, the number of unique subjects described in the literature is likely approximately 1039. We conclude that approximately one of every two individuals described in a psychedelics fMRI study has been described in at least one additional study. This observation was most striking for LSD, where the average ratio of unique individuals studied per publication was only five participants. Specifically, despite the fact that we reviewed 17 papers examining LSD with fMRI, it appears that these data only consider 85 unique individuals (Table 1).

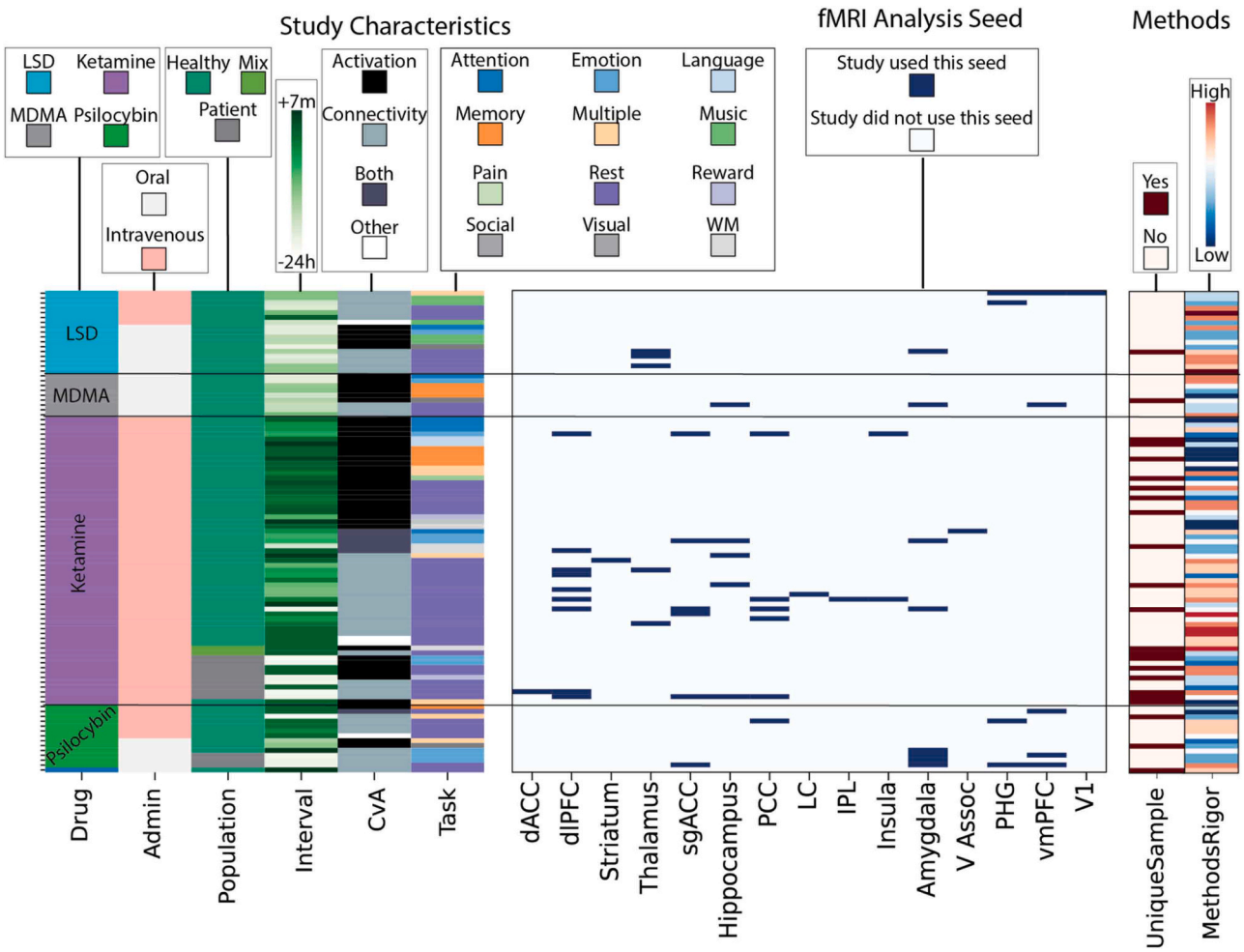

Fig. 2. A visual representation of variability across psychedelics fMRI studies. The leftmost matrix represents aspects of study design, the center matrix represents seed region in the case of seed-based studies, and the rightmost matrix represents key methodological considerations. In the leftmost matrix, a subset of vectors represent variation in key aspects of study and analytic designs across psychedelics studies. Each row represents a study, and rows are sorted by (in order) drug, method of administration, dose, study population, interval, connectivity vs activation (CvA), and task. For the "interval" column, all intervals were rank ordered for ease of visualization. For the "Dose" column, dosages were grouped within-drug based on identical/similar dosages, coded by group, ordered by dosage, and then min-max normalized within-drug. Note that the "Unique Sample" column indicates whether the sample used in this study was entirely unique and not used by another study. This image indicates a high degree of dissimilarity in analytic and overall study design across studies, with a great deal of data recycling and variable methodological rigor (Admin = Route of administration; CvA = Connectivity vs Activation; dACC = dorsal anterior cingulate cortes; dlPFC = dorsolateral prefrontal cortex; IPL = inferior parietal lobe; LC = locus coeruleus PCC = posterior cingulate cortex; sgACC = subgenual anterior cingulate cortex; V Assoc = Visual association cortex).

Table 1
Summary of unique samples across psychedelics fMRI studies.

Drug	Total Pubs			Unique Samples			Total N			Total Pubs: Total N		
	Total	**Task**	**Rest**	**Total**	**Task**	**Rest**	**Total**	**Task**	**Rest**	**Total**	**Task**	**Rest**
Ketamine	**59**	27	32	**37**	21	22	**753**	404	472	**12.8**	15.0	14.8
LSD	**17**	9	8	**4**	3	4	**85**	64	80	**5.0**	7.1	10.0
MDMA	**8**	6	3	**4**	4	1	**80**	72	40	**10.0**	12.0	13.3
Psilocybin	**14**	7	7	**6**	5	3	**141**	113	45	**10.1**	16.1	6.4

3.3. Heterogeneity in task-based fMRI studies of psychedelics

Another major source of heterogeneity in the studies included in our search was the large variety of fMRI task paradigms used (Fig. 2). This feature is best exemplified by the MDMA literature, for which 8 publications described findings across five different types of tasks. Although five different LSD studies used a music-based task (though these were based on only two different samples), other studies engaged in attention, emotion, and social cognition tasks. Three psilocybin studies (from two samples) used emotional cognition tasks; other studies used tasks focused on learning and memory and social cognition. There was also a great variability in tasks across ketamine studies: five studies (from four samples) used emotion tasks, four studies (from three samples) used attention tasks, four studies (from four samples) used learning/memory tasks, four studies (from four samples) used working memory tasks, two studies (from two samples) used reward-based tasks, two studies (from two samples) used language tasks, and one each used pain and visual tasks. It is important to note here that these task categories are broad

categorizations; very few studies used identical tasks, and occasionally tasks within the same category were designed to evaluate different cognitive mechanisms. This variety of study design ultimately poses an obstacle in integrating information across studies.

3.4. Existing studies often do not adequately control for Type I error

For each paper included, reviewers evaluated whether contemporary procedures for control of multiple comparisons were present. Notably, for each drug included in the review, only slightly over half of included studies reported use of methods for adequate Type I error control (Fig. 3**B,** Table S1). Specifically, 41% of ketamine studies, 43% of LSD studies, 46% of psilocybin studies and 50% of MDMA studies did not meet our (liberal) criteria for adequate Type 1 error control.

3.5. Motion artifact is often not considered in fMRI studies of psychedelics

Next, we examined each publication and assigned a numerical score (0−2) to denote whether publications provided information about a) whether studies described exclusion of participants for in-scanner motion and b) whether studies evaluated or controlled for the impact of motion artifact in hypothesis testing (Table S1). Across all four drugs surveyed, 61% of included publications fulfilled at least one criterion for recommended motion correction strategies (Fig. 3**A**). Within each drug type, papers examining LSD and MDMA had the highest percentage of publications that transparently addressed motion artifact (94% and 75%, respectively). In contrast, 49% of ketamine publications and 62% of psilocybin publications appeared to account for the impact of motion in their analyses. Meanwhile, only 15% of both ketamine and psilocybin studies fulfilled both criteria for motion correction strategies; 41% of LSD studies and 62% of MDMA studies fulfilled both criteria.

3.6. Methodological rigor in fMRI studies of psychedelics has improved over time

The above results indicate inconsistent adherence to contemporary methodological standards for control of Type I error and the confounding influence of motion artifact. However, when an aggregated methodological rigor score was plotted versus publication date, we observed that rigor has improved over time (Fig. 3**C**). The overall quality of methodological rigor improved substantially between the turn of the century and the mid 2010 s, after which it appears to have plateaued. The amount of variation in study quality has also increased over time, with substantial ongoing methodological heterogeneity being present even among relatively recent studies.

3.7. Task-based studies tend to employ less rigorous Type I error correction

Assessment of strategies to overcome motion artifact and Type I error correction were further stratified by whether studies used resting-state or task-based paradigms. Qualitatively, there were no strong differences between paradigms in strategies to overcome motion artifact (Fig. 3**D**), but task-based studies were less likely to employ satisfactory Type I error correction across all drugs except MDMA (Fig. 3**E**). However, the same upward trend (followed by plateau) of methodological

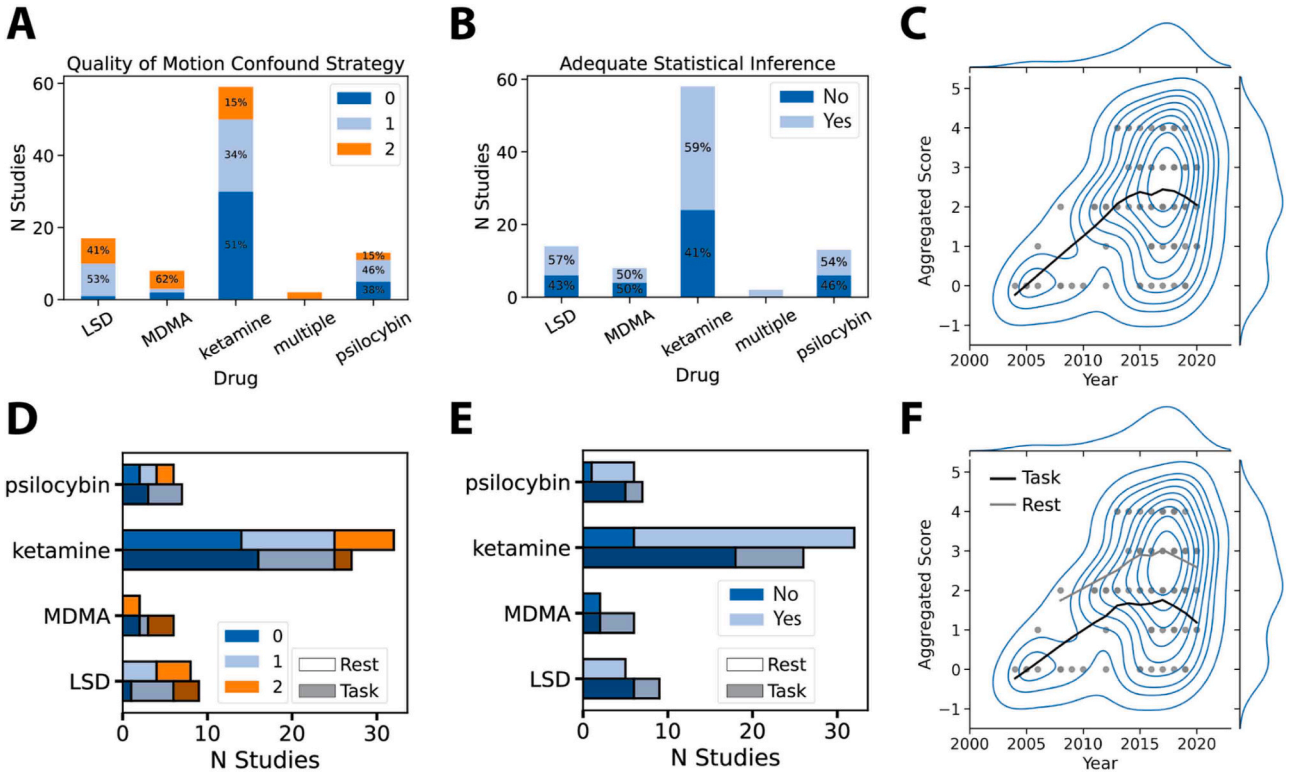

Fig. 3. Summary of methodological rigor indices across studies. A) Studies were scored from 0 to 2 based on the degree to which they controlled for the confounding effect of in-scanner head motion on the BOLD signal (see Methods). One point was given if studies removed scans due to motion, and another if they evaluated motion as part of hypothesis testing. Contemporary studies are encouraged to perform both of these strategies, but nearly 40% of studies reviewed engaged in neither strategy and only 25% engaged in both. **B)** Studies were given a binary score indicating whether they used unbiased brain masks and applied adequate correction against Type I error when assessing the main effect of treatment. Across different drugs, 40–50% of studies did not fulfill these criteria. **C)** The motion and Type I error scores were combined into an aggregated score (the binary Type I error was multiplied by two to achieve equal weighting between it and motion correction score). A positive trend was observed in improving methodological rigor between the early 2000 s and mid 2010 s. However, this trend leveled off at a score of 3, and a great deal of variability in methodological rigor among studies persisted even up to 2020. **D) - F)** The same as A) - C) but further stratified by whether studies were task-based (Task, darker colors) or resting-state (Rest, lighter colors).

rigor over time was seen across both paradigms (Fig. 3**F**).

3.8. Selection of studies for further qualitative review and discussion

Based on the methodological review conducted here, we next identified a subset of 40 publications that had adequate control of Type I error and any address of motion artifact. These included 25 ketamine papers, eight LSD papers, five psilocybin papers, and two MDMA papers. Note that many of these publications did have overlapping samples, which we accounted for in our qualitative synthesis and discussion of this literature (see below).

3.9. Consistent findings across studies administering classic psychedelics

Following our methodological review, we sought to integrate coherent findings across published studies that met contemporary standards for Type I error control and reporting of motion artifact. In total, this included five psilocybin studies and eight LSD studies. Below, we discuss the (13 total) studies of LSD and psilocybin jointly as they are classic psychedelics that are thought to act on similar receptors. Supporting this notion, many studies reported a brainwide correlation between psilocybin and LSD effects on brain connectivity (Fig. 4**A,B**) (Tagliazucchi et al., 2016; Preller et al., 2020, 2018a). Three of the 13 studies that met criteria for review are not discussed below – Carhart-Harris et al. (2017) (Carhart-Harris et al., 2017) did not report any significant findings for the contrast of interest, while (Schmidt et al., 2018a) and (Preller et al., 2018b) used attention tasks while all other studies used resting-state protocols.

Some consistency was seen across the ten LSD/psilocybin studies reviewed, particularly among studies conducting network-based or whole-brain connectivity analyses. Several LSD (Tagliazucchi et al., 2016; Preller et al., 2018a; Müller et al., 2018) and psilocybin (Lord et al., 2019) studies reported an increase in global brain connectivity, most notably in the association cortex (Fig. 4**D,H**). Studying LSD in two different samples, Preller et al. (2018a) and Tagliazucchi et al. (2016) both reported a global increase in connectivity across association cortex. Related, Preller et al. reported a global decrease in connectivity across sensory cortex, while Müller et al. (2018) (same cohort as Tagliazucchi et al.) also reported increased connectivity between association cortex networks. Echoing these results under psilocybin administration, Roseman et al. (2014) reported a strong increase in connectivity between association cortex networks, but a decrease in connectivity between certain sensory cortex networks (Fig. 4**E**). Applying time-varying connectivity analysis to study phase coherence of brain networks within the same cohort, Lord et al. (2019) reported that psilocybin administration results in a more frequent occurrence of an unpatterned state of global (rather than network-specific) connectivity (Fig. 4**G**). Increased connectivity among association cortex was therefore consistently reported across studies, as was a finding of opposing decrease in sensory cortex connectivity. Some of the above studies also noted a decrease in modularity (Fig. 4**C**) and/or a decrease in within-network connectivity (Tagliazucchi et al., 2016; Müller et al., 2018; Roseman et al., 2014), though these conclusions were reached using only one sample each for LSD and psilocybin. These three studies reporting decreases in modularity specifically noted a loss of segregation that included greater connectivity of sensorimotor cortex with association cortex.

A few other fairly consistent findings emerged from this literature. Increased thalamic and general subcortical connectivity with association cortex was reported consistently across LSD and psychedelics studies (Fig. 4**F**) (Preller et al., 2018a; Müller et al., 2018; Mueller et al., 2017; Bershad et al., 2020; Lebedev et al., 2015). Furthermore, two different LSD studies and one psilocybin study found a strong overlap between 5HT2R distribution and drug-induced change to global brain connectivity (Tagliazucchi et al., 2016; Preller et al., 2018a), a finding supported by the partial abolition of effects when LSD was co-administered with the 5HT2R agonist ketanserin (Preller et al., 2018a). This finding is not unexpected given that both LSD and psilocybin act principally on 5HT2R receptors.

A caveat to the above consistency across studies is that they may depend on a specific methodological step: the use of global signal regression in de-noising to control for brain-wide fluctuations that often arise from head motion and physiological artifact. Preller et al (Preller et al., 2018a). noted that regression of global BOLD signal led to complete reversal of results: when global signal regression was employed, LSD administration led to decreased connectivity among association cortex and increased connectivity in sensory cortex (Fig. 4**H,I**). This finding was closely reproduced in a follow-up study from the same group, this time when evaluating psilocybin administration in a new sample (Preller et al., 2020) following global signal regression. Notably, all studies previously discussed in this section either did not use global signal regression or did not report its use.

3.10. Consistent findings across studies administering ketamine

Twenty-three ketamine studies fulfilled our criteria for methodological rigor and underwent further review. Among these studies were several that administered ketamine mid-scan and examined BOLD changes time-locked to drug administration. These studies resulted in a highly consistent finding, namely that ketamine generally increases BOLD activation, a convergent finding across seven different samples (Javitt et al., 2018; De Simoni et al., 2013; Driesen et al., 2013; McMillan et al., 2020; Downey et al., 2016; Doyle et al., 2013; Höflich et al., 2017). Moreover, multiple studies found peak activation to localize to the dorsolateral prefrontal cortex, particularly to a medial region just dorsal to the caudal anterior cingulate (Javitt et al., 2018; De Simoni et al., 2013; McMillan et al., 2020; Downey et al., 2016; Höflich et al., 2017) (Fig. 5**A**). Importantly, this effect remained robust to regression of several different motion and physiological parameters (De Simoni et al., 2013; McMillan et al., 2020).

A few other consistent findings emerged from the ketamine literature, though these findings were each supported by only a few studies each. Subcortical activation was seen across multiple studies, as was increased within-subcortex connectivity (Joules et al., 2015; Höflich et al., 2015). Despite the increase in activation associated with ketamine administration, several studies reported a decrease in activation specifically in the sgACC (De Simoni et al., 2013; McMillan et al., 2020; Doyle et al., 2013; Wong et al., 2016), a region implicated in major depressive disorder (Drevets et al., 2008). However, two independent studies showed that this finding was abolished after correction for motion and/or physiological confounds (De Simoni et al., 2013; McMillan et al., 2020). In addition, only two of the 23 ketamine studies involved tasks other than rest, but both showed reduced activation of task-positive networks (D'Souza et al., 2018; Steffens et al., 2016). The remaining studies reviewed offered little in the way of reproducible or consistent results – most were connectivity studies and varied in both the direction and location of alterations reported in response to ketamine.

3.11. MDMA literature is nascent

Of eight BOLD imaging studies of participants administered MDMA, only two (Schmidt et al., 2018b, 2017) met our criteria for methodological rigor. Both studies were published by the same group, examining the effects of MDMA on two different tasks in the same sample of participants. One study (Schmidt et al., 2018b) found that MDMA increased activation of frontal-parietal regions during a Go/No-go task, while the other study (Schmidt et al., 2017) found no differences in activation compared to placebo in response to viewing fearful faces.

4. Discussion

In this review, we sought to synthesize functional imaging studies of

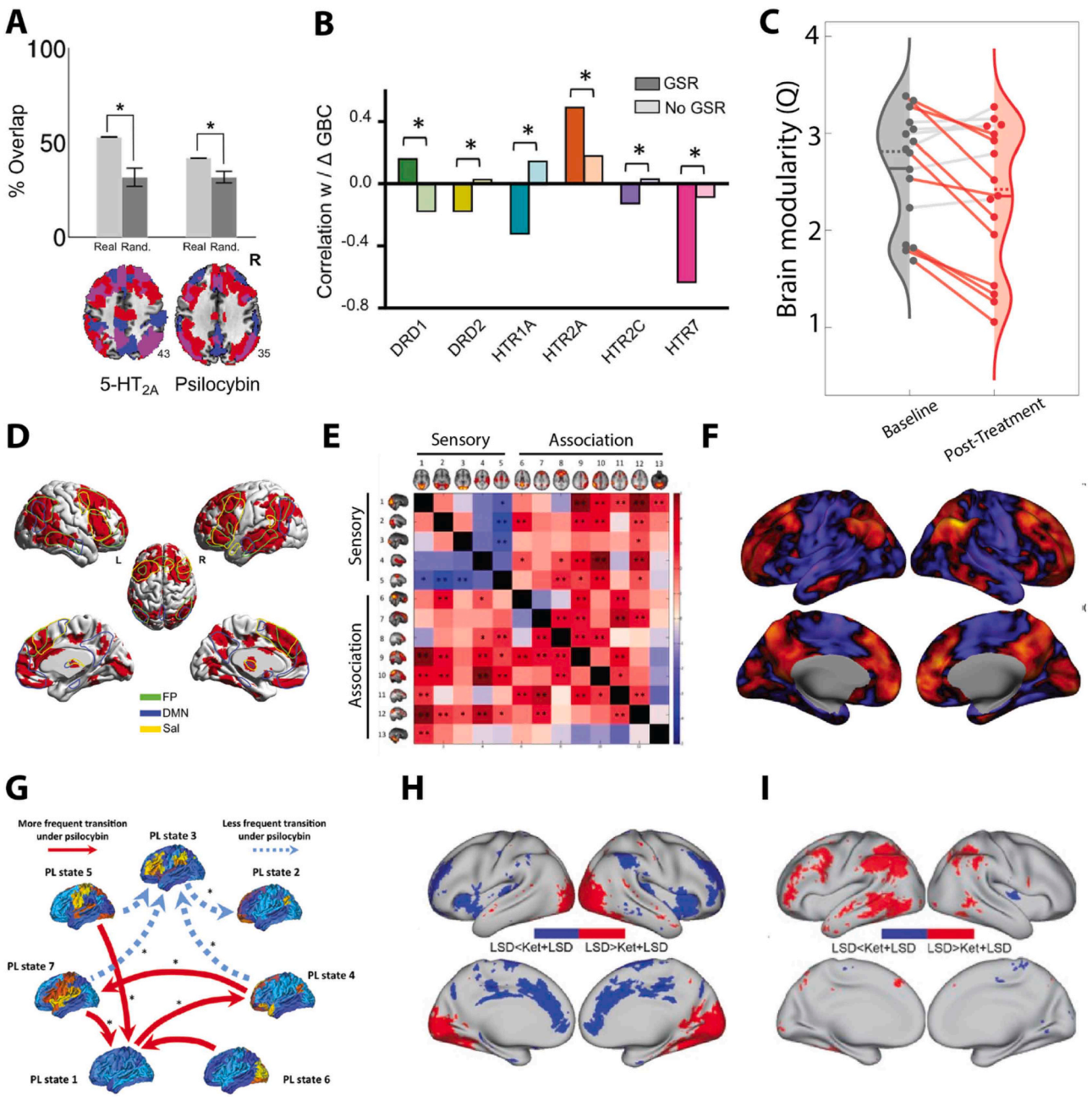

Fig. 4. Convergent changes in functional connectivity following acute administration of classic psychedelics. Exemplar figures were collected to display convergent findings across studies. **A**) Increases in global brain connectivity related to LSD administration show significant overlap with 5-HT2A receptor distribution. % overlap indicates spatial overlap between thresholded 5HT2A receptor map and *t*-map of increased connectivity after psilocybin administration, compared to spatial permutations. **B)** Whole-brain correlations between various cortical neurotransmitter maps and change in global brain connectivity associated with LSD. **C)** An example study showing psilocybin administration to reduce brain network modularity. **D)** An example study showing increased connectivity between association cortex with the rest of the brain under LSD administration compared to placebo (shown in red). FP = frontoparietal network; DMN = default mode network; Sal = Salience network. **E)** An example study showing decreased connectivity (cool colors) between sensory cortex network, but increased connectivity (warm colors) between sensory and association cortex networks, under psilocybin administration. **F)** An example study showing increased connectivity between subcortex and association cortex under LSD administration. The current example is unthresholded (several regions do survive FDR correction) and visualizes changes in connectivity with the amygdala, but other studies have also described similar findings seeding the thalamus. **G)** One study showing psilocybin administration leads to an increased transition into a state of global connectivity and away from states of local connectivity. **H,I)** FDR corrected findings of regional changes in global connectivity after LSD administration (compared to administration of LSD plus 5HTR2A agonist ketanserin) without (**H**) and with (**I**) correction for global signal. See paper for uncorrected results.

(a) Adapted from Tagliazucchi et al., 2016. (b) Adapted from Preller et al., 2018a. (c) Adapted from Daws et al., 2022. (d) Adapted from Tagliazucchi et al., 2016. (e) Adapted from Roseman et al., 2014. (f) Adapted from Bershad et al., 2020. (g) Adapted from Lord et al., 2019. (h) and (i) Adapted from Preller et al., 2018a.

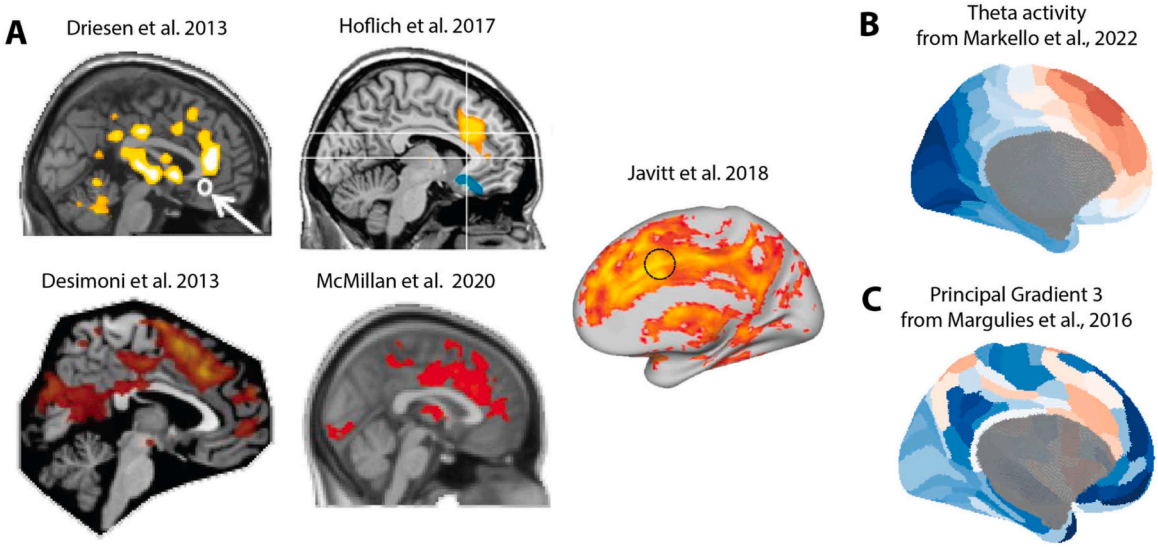

Fig. 5. Convergent changes in BOLD signal following ketamine administration across studies. A) Five different within-subject studies measuring BOLD changes time-locked to ketamine administration show increased BOLD signal in the dorsomedial prefrontal cortex and/or caudal anterior cingulate cortex. Other regions of convergent BOLD activation not shown include insula, dorsolateral prefrontal cortex and basal ganglia. **B)** Cortical rendering of the spatial topography of theta band activity as measured through magnetoencephalography (maps created using data from Markello et al., 2022 *Nature Methods*). **C)** Cortical rendering of the third principal gradient of functional connectivity (created using data from Margulies et al., 2016 *PNAS*).

the acute administration of psychedelic agents. Our goals were two-fold. First, we sought to systematically examine important methodological characteristics of these studies, including study design, sample size, sample uniqueness, handling of in-scanner motion, and Type I error control. Second, we aimed to qualitatively review a sub-sample of studies (N = 40) that used contemporary methods for control of motion artifact and Type I error. Below, after discussing the current methodological landscape, we qualitatively integrate these findings.

4.1. Functional imaging studies of acute responses to psychedelics are notable for heterogeneity in study design, sample size, and methodological rigor

Our review revealed that fMRI studies of acute responses to psychedelics employ widely varying study designs, often studied small (and frequently overlapping) samples, and have markedly heterogeneous adherence to best practices regarding control of Type I error and motion artifact. Many studies reviewed used strong overall designs, with double-blind, cross-over designs being the most prevalent, sometimes including additional conditions with pharmacological blockade of target receptors. However, the timing and dosing regimen of drug administration were highly inconsistent across studies. As several studies reported that the distribution and magnitude of effects on BOLD signal varied depending on timing and dosage (Chen et al., 2019; Javitt et al., 2018; Li et al., 2020; Preller et al., 2020), this identifies a critical issue. Concerted efforts to understand the pharmacological effects of psychedelics on brain function would likely benefit from standardization across studies, as well as more studies that carefully parse how dose and timing impact fMRI measurements. Improved standardization efforts could also be extended to the selection of in-scanner tasks across psychedelics studies. The great variability in tasks we observed is probably due to the diverse hypotheses regarding the impact of psychedelics on brain function and behavior. However, such diversity also makes it challenging to establish reproducible findings across studies, or even across studies probing similar cognitive constructs. We also found task-based studies were less likely to deploy adequate Type I statistical error correction, perhaps betraying a tendency toward smaller effect sizes. Moving forward, integration of a resting-state acquisition into scanning protocols – which serves as a simple and easily reproducible protocol – could help facilitate data pooling and replication efforts.

Variability in study design was accompanied by use of relatively small and often overlapping samples. Although several larger studies (notably of ketamine) have been conducted more recently, the average sample size of within-subject designs was under 20 participants, and the average sample of between-subject designs included fewer than 30 participants. While small samples were the rule rather than exception, there is a positive trend towards increasingly large samples over the date ranges we evaluated in this review. However, it should be noted that many of the studies reviewed studied the same participants. Sometimes, the same individuals underwent different tasks as part of a larger study protocol, and sometimes the same data were simply re-analyzed. Given the high level of difficulty of this research, including massive regulatory overhead and the historically sparse funding, this phenomenon is not surprising. However, it should be noted that reporting of previously published samples was inconsistent across the studies reviewed. In 25% of cases, samples shared across studies were deduced through careful review of sample sizes, sample characteristics and study protocols, rather than through explicit reporting. When combined with small samples and accumulating literature suggesting functional networks are person-specific (Finn et al., 2015; Dickie et al., 2018; Cui et al., 2020), such practices have the potential to substantially reduce the generalizability of reported findings. Though not quantified here, study recruitment was generally enriched with people with prior psychedelic experience, which may further reduce the generalizability of results. Additionally, it should be noted that studies of psychedelics are often quite limited in the diversity of participant demographics, which may further limit generalizability to the broader population (Thrul and Garcia-Romeu, 2021; Michaels et al., 2018; George et al., 2019).

Another obstacle to generalizable findings apparent in the reviewed literature was the heterogeneous application of contemporary standards for Type I error control and accounting for motion artifact. Especially when combined with small samples, this limitation raises the possibility that many of the reported results in the literature are, in fact, false positives or directly result from systematic confounding influence of motion across study groups or conditions. Though the field of psychedelics fMRI literature is young, examples of these concerns affecting the course of research have already been noted. For example, although early studies showed a reduction of sgACC activity upon ketamine

administration (Deakin et al., 2008), that finding was later found to be explained by motion (McMillan et al., 2020), but not before inspiring follow-up studies to focus on the mPFC as a seed for connectivity analysis (De Simoni et al., 2013; Wong et al., 2016; Gärtner et al., 2019; Scheidegger et al., 2012). This example is particularly disconcerting considering that a similar finding of decreased sgACC activity was reported for psilocybin (Carhart-Harris et al., 2012), a drug that has undergone less methodological scrutiny in the literature. These issues go beyond influencing scientific studies – media coverage of such findings often assumes they are vetted for scientific methodology through peer review, raising the likelihood that misinterpreted data is communicated to wide audiences. Fortunately, concerns with Type I error correction can be easily remediated (Poldrack et al., 2017; Roiser et al., 2016). While the methods for denoising continue to rapidly evolve (Ciric et al., 2018), transparent reporting of differences in motion across study conditions – combined with sensitivity analyses when warranted – would bolster confidence in reported findings.

4.2. Classic psychedelics influence BOLD changes along the sensorimotor-association axis of brain organization

Among selected studies that met certain contemporary standards for motion correction and type I error correction, consistent findings emerged. Both LSD and psilocybin frequently reported increased whole-brain connectivity, but particularly in association cortex and between sensory and association cortex. Another frequent finding was a general decrease in network modularity after administration of these two drugs. Decreased network modularity indicates less segregation of distinct networks (i.e., usually distinct networks becoming more functionally coupled), which is in line with findings of increased whole-brain connectivity. These findings have been supported by psychedelics studies that were published more recently (after our systematic review was conducted). Girn et al (Girn et al., 2022) used an LSD sample that had already been published but similarly found decreased modularity accompanying LSD administration in this dataset. Meanwhile, Daws et al. reported the same phenomenon in a new psilocybin sample (Daws et al., 2022).

Together, these findings suggest that classic psychedelics have differential effects on sensorimotor and association cortices, perhaps disrupting the normative connectivity dynamics between these types of cortex. Accordingly, it appears that the acute effects of psychedelics on fMRI-derived measures may vary across the brain's sensorimotor-association (S-A) axis. The S-A axis is the principal spatial dimension of variation in cortical connectivity (Margulies et al., 2016), is one of three spatiotemporal components explaining group-level BOLD variability (Bolt et al., 2022), and more broadly is a major dimension of variability across many different brain properties (Sydnor et al., 2021). Several other lines of evidence link the S-A axis to brain changes in response to psychedelics. The thalamus is understood to developmentally regulate the identity of sensory and association areas, and to facilitate control of cortical dynamics across the S-A axis (Anderson et al., 2018) via thalamocortical connections (Vue et al., 2013; O'Leary et al., 2007; Müller et al., 2020). A consistent finding across studies reviewed was that thalamic and general subcortical connectivity increased with association cortex upon psychedelic administration. Intriguely, several studies reviewed also found the pattern of connectivity increase after psychedelic administration overlapped with the distribution of 5HT2a receptors in the brain. Such findings are related to the S-A axis: recent work has described an axis of receptor distribution variation in the brain that is anchored on one side by 5HT2a receptors, is strongly correlated ($r = 0.9$) with the S-A axis (Luppi et al., 2022), and captures variation in the effects of mind-altering drugs on fMRI connectivity. These findings together emphasize that expression levels of 5HT2R across the S-A axis may in part explain the diverging effects of classic psychedelics across the brain.

Based on these findings, one promising avenue for future research is to explore how psychedelics affect large-scale properties of brain organization, such as the S-A axis. This idea is explored in a recent study examining the effects of LSD on the S-A axis at the individual level (Girn et al., 2022). This study found that the S-A axis was disrupted, which was explained by an increased integration – i.e., a loss of typical network segregation – between unimodal (e.g. sensory) and transmodal (e.g. association) areas. The study also found that psychedelics perturb other major axes of functional activity that differentiate sensory and executive systems. Importantly, these findings potentially nominate classic psychedelics as pharmacological tools to experimentally manipulate whole-brain functional organization. However, it should be noted that the diminished network segregation observed in response to psychedelics is typically thought to be a negative marker of brain health in other contexts: network segregation *increases* as part of healthy development in youth (Bassett et al., 2018), is associated with better cognitive performance (Keller et al., 2023), declines in aging (Chan et al., 2014), and is degraded in neurodegenerative and neuropsychiatric conditions (Chan et al., 2021). Thus, the acute impact of psychedelics on functional brain organization cannot be straightforwardly interpreted as therapeutic or neuroprotective.

One thread emerging from the classic psychedelics literature that merits further investigation is the apparent inversion of results that occurs when global signal is regressed out from the data before analysis. This reversal of findings suggest that LSD and psilocybin may alter global signal in the brain, which itself leads to alterations in brain network organization. In the studies that reported this finding (Preller et al., 2020, 2018a), the influence of global signal on network topology was apparently not caused by changes to signal variance or amplitude, but rather by certain regions contributing differentially to the global signal during drug administration. Notably, recent research (Power et al., 2017; Gratton et al., 2020; Lynch et al., 2020) suggests that changes in respiration rate and pattern may dramatically impact the global signal, and that global signal regression may remove substantial respiratory variance. Given the potential for psychedelic administration to alter respiratory patterns – i.e., induce deep breathing – future studies should evaluate such physiologic data carefully alongside preprocessing choices, such as global signal regression.

4.3. Ketamine increases in BOLD activation especially localized to dorsomedial prefrontal cortex

Among ketamine administration studies selected for qualitative review, several consistent findings emerged. First, ketamine generally increased whole-brain activation. Second, this effect was especially strong in the dorsomedial prefrontal cortex. Further studies will be necessary to understand how these observations coalesce with the current understanding of ketamine mechanisms of action. Ketamine is an NMDA receptor antagonist and is thought to preferentially block NMDA receptors on GABAergic interneurons, therefore leading to reduced inhibition and increased activation of pyramidal neurons and increased release of glutamate (Duman et al., 2019; Gerhard et al., 2020; Luscher et al., 2020). The overall effect leads to higher activity in cortical circuits, which seems to concur with the consistent findings of ketamine increasing overall BOLD signal at rest. However, future work will be needed to determine how this mechanism relates to the particular pattern of BOLD activation consistently observed across studies, and how this might be relevant to behavioral effects of ketamine administration.

Intriguingly, results in the dorsomedial prefrontal cortex appeared again in another study, which reported increased fractional amplitude of low-frequency fluctuations (fALFF) in this region, and decreased connectivity to the posterior cingulate (Li et al., 2020). While unrelated to BOLD, McMillan et al (McMillan et al., 2020) noted ketamine also increased short-wavelength frontoventral-sourced MEG signals, such as theta rhythms. This is interesting given that the apparent cortical source of theta oscillations as measured with MEG (McMillan et al., 2020;

Markello et al., 2022) is a region overlapping with the dorsal PFC activation peak seen after ketamine administration (Fig. 5**B**). Importantly, this region is also a critical component of the salience network, along with other subcortical structures. Perhaps relevant, subcortical activation was seen across multiple studies reviewed here, again implicating the salience network. Finally, the dorsomedial frontal cortex is a region situated at one pole of the third principal gradient of functional connectivity (Margulies et al., 2016) (Fig. 5**C**), which is thought to differentiate executive or task-positive systems from the rest of the cortex (Girn et al., 2022). This is interesting given that the two task-based ketamine administration studies we reviewed both showed reduced activation of task-positive networks. These findings are generated from a very small samples of studies, but warrant further investigation.

4.4. Limitations

This systematic review has several potential limitations. First, many aspects of study design and data analysis were not considered, particularly relating to data preprocessing. Preprocessing strategies and pipelines beyond those discussed here can have a tremendous effect on results of fMRI analysis (Botvinik-Nezer et al., 2020), and future studies should catalog this important aspect of study variability. Second, we could not ascertain with full confidence which papers analyzed data from the same participant sample. Thus some of our estimates could be incorrect. We tried to be transparent about this limitation by documenting our confidence in the assignment of shared samples. Third, while we expanded upon other reviews (McCulloch et al., 2021) by including task-based studies and atypical psychedelics, other drugs were not reviewed. For example, we chose not to review studies administering cannabis, and while we did identify a few fMRI studies of DMT (ayahuasca), there were too few to warrant review. Cannabis was excluded from this meta-analysis for several reasons: A) cannabis would be the only drug examined that acts on the cannabinoid system, B) it has been covered extensively in other reviews, C) cannabis is typically not considered a psychedelic except in very high doses. Fourth, as stipulated in our pre-registered meta-analysis, studies with both healthy control and patient populations were included in our systematic review, which has the potential to increase heterogeneity among studies. However, such heterogeneity did not seem to impact our qualitative synthesis or our conclusions: of three studies with patient populations selected for qualitative review, two showed results highly consistent with studies with control populations (Downey et al., 2016; D'Souza et al., 2018), while the last reported no significant findings for the contrast of interest (Carhart-Harris et al., 2017). Fifth, we focused on studies that examined acute fMRI responses to psychedelic administration, ignoring brain-behavior relationships reported in some studies. Sixth, despite concerted assessment of methodological rigor, we tended to err on the side of leniency when scoring manuscripts on their adherence to current standards. Therefore, despite our structured assessment, we may have overestimated the rigor of certain studies. Notably, as detailed pre-registration of analyses was uncommon, we were unable to assess the potential impact of overfitting or repeated multiple testing ("p-hacking") in this literature.

5. Conclusion

The last twenty years has seen a rapid increase in studies examining the pharmacological, cognitive and therapeutic effects of psychedelics. Driven by a more tractable regulatory environment and by promising clinical data, use of fMRI in translational studies that seek to find non-invasive correlates of acute functional responses to psychedelics has increased. The present review emphasizes both the substantial pitfalls and great promise of this research. As in many other rapidly-moving subfields of functional neuroimaging research, there is a risk of forgetting – and then painfully re-learning – many of the methodological lessons of the past. Moving forward, two of the current recommendations for best practices in fMRI research in other domains could be easily applied to studies of psychedelics. First and foremost is clear pre-registration of hypotheses and outcomes alongside detailed analytic plans. Although clinical trials often include many of these elements, functional imaging details are often specified only vaguely, leaving room for substantial analytic flexibility and manual overfitting of data. Second, use and open sharing of highly reproducible analytic methods, such as containerized preprocessing pipelines and analytic notebooks, will likely increase confidence in any results obtained. Though not discussed at length here, researcher degrees of freedom and p-hacking have been at the center of the reproducibility crisis in neuroscience (Botvinik-Nezer et al., 2020; Head et al., 2015), and are of particular concern in small-sample, "hot-topic", commercializable research. Given the broad excitement in psychedelic research and the emergence of convergent findings in the literature, significant investment in reproducible practices is especially important. Indeed, consistent findings reviewed here – including classic psychedelics affecting connectivity structure of S-A axis and serotonergic cortex and of ketamine increasing activation of dorsomedial frontal cortex – provide hope for future research. Together, fMRI studies of psychedelics have the potential to significantly increase our understanding of both mechanisms of psychedelic action and in stratifying patients for treatment with these powerful agents.

Declaration of Competing Interest

Robert H. Dworkin, PhD, has received in the past 5 years research grants and contracts from the US Food and Drug Administration and the US National Institutes of Health, and compensation for serving on advisory boards or consulting on clinical trial methods from Abide, Acadia, Adynxx, Analgesic Solutions, Aptinyx, Aquinox, Asahi Kasei, Astellas, Beckley, Biogen, Biohaven, Biosplice, Boston Scientific, Braeburn, Cardialen, Centrexion, Chiesi, Chromocell, Clexio, Collegium, CoimbiGene, Confo, Decibel, Editas, Eli Lilly, Endo, Ethismos (equity), Eupraxia, Exicure, GlaxoSmithKline, Glenmark, Gloriana, Hope, Juca, Kriya, Lotus, Mainstay, Merck, Mind Medicine (also equity), Neumentum, Neurana, NeuroBo, Novaremed, Novartis, OliPass, Orion, Oxford Cannabinoid Technologies, Pfizer, Q-State, Reckitt Benckiser, Regenacy (also equity), Rho, Sangamo, Sanifit, Scilex, Semnur, SIMR Biotech, Sinfonia, SK Biopharmaceuticals, Sollis, SPM Therapeutics, SPRIM Health, Teva, Theranexus, Vertex, Vizuri, and WCG.

Over the last three years, Eric Strain has had grant funding to his institution from NIH; editing payments from: Wolters-Kluwer, Elsevier (ended); consulting fees from: Pear Therapeutics (ended), Fast Track Drugs and Biologics, Otsuka Pharmaceutical Development and Commercialization, Inc., Cerevel; honoraria for talks: Rutgers University; medical devices supplied to his institution for his research: Masimo; and medical-legal consultations. He also has served on the board of directors (unpaid) for a treatment program: Ashley Addiction Treatment.

Data Availability

No data was used for the research described in the article.

Acknowledgements

This work was supported by Analgesic, Anesthetic, and Addiction Clinical Trial Translations, Innovations, Opportunities, and Network (ACTTION). Author JWV acknowledges support from the SciLifeLab & Wallenberg Data Driven Life Science Program (grant: KAW 2020.0239).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.neubiorev.2023.105421.

References

Anderson, K.M., et al., 2018. Gene expression links functional networks across cortex and striatum. Nat. Commun. 9, 1428.

Bassett, D.S., Xia, C.H., Satterthwaite, T.D., 2018. Understanding the emergence of neuropsychiatric disorders with network neuroscience. Biol. Psychiatry Cogn. Neurosci. Neuroimaging 3, 742–753.

Bennett, C.M., Miller, M.B., Wolford, G.L., 2009. Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: An argument for multiple comparisons correction. Neuroimage 47, S125.

Bershad, A.K., et al., 2020. Preliminary report on the effects of a low dose of LSD on resting-state amygdala functional connectivity. Biol. Psychiatry Cogn. Neurosci. Neuroimaging 5, 461–467.

Bolt, T., et al., 2022. A parsimonious description of global functional brain organization in three spatiotemporal patterns. Nat. Neurosci. 25, 1093–1103.

Botvinik-Nezer, R., et al., 2020. Variability in the analysis of a single neuroimaging dataset by many teams. Nature 582, 84–88.

Carhart-Harris, R., et al., 2021. Trial of psilocybin versus escitalopram for depression. New Engl. J. Med. 384, 1402–1411.

Carhart-Harris, R.L., et al., 2012. Neural correlates of the psychedelic state as determined by fMRI studies with psilocybin. Proc. Natl. Acad. Sci. USA 109, 2138–2143.

Carhart-Harris, R.L., et al., 2017. Psilocybin for treatment-resistant depression: fMRI-measured brain mechanisms. Sci. Rep. 7, 1–11.

Carmichael, O., et al., 2018. The role of fMRI in drug development. Drug Discov. Today 23, 333–348.

Carter, C.S., Lesh, T.A., Barch, D.M., 2016. Thresholds, power, and sample sizes in clinical neuroimaging (Preprint at). Biol. Psychiatry.: Cogn. Neurosci. Neuroimaging vol. 1, 99–100. https://doi.org/10.1016/j.bpsc.2016.01.005.

Chan, M.Y., et al., 2021. Long-term prognosis and educational determinants of brain network decline in older adult individuals. Nat. Aging 1, 1053–1067.

Chan, M.Y., Park, D.C., Savalia, N.K., Petersen, S.E., Wig, G.S., 2014. Decreased segregation of brain systems across the healthy adult lifespan. Proc. Natl. Acad. Sci. USA 111, E4997–5006.

Chen, M.-H., et al., 2019. Antidepressant and antisuicidal effects of ketamine on the functional connectivity of prefrontal cortex-related circuits in treatment-resistant depression: a double-blind, placebo-controlled, randomized, longitudinal resting fMRI study. J. Affect. Disord. 259, 15–20.

Ciric, R., et al., 2017. Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. Neuroimage 154, 174–187.

Ciric, R., et al., 2018. Mitigating head motion artifact in functional connectivity MRI. Nat. Protoc. 13, 2801–2826.

Cui, Z., et al., 2020. Individual variation in functional topography of association networks in youth. Neuron 106 (340–353), e8.

D'Souza, D.C., et al., 2018. Dose-related target occupancy and effects on circuitry, behavior, and neuroplasticity of the glycine transporter-1 inhibitor PF-03463275 in healthy and schizophrenia subjects. Biol. Psychiatry 84, 413–421.

Daws, R.E., et al., 2022. Increased global integration in the brain after psilocybin therapy for depression. Nat. Med. 28, 844–851.

De Simoni, S., et al., 2013. Test-retest reliability of the BOLD pharmacological MRI response to ketamine in healthy volunteers. Neuroimage 64, 75–90.

Deakin, J.F.W., et al., 2008. Glutamate and the neural basis of the subjective effects of ketamine: a pharmaco-magnetic resonance imaging study. Arch. Gen. Psychiatry 65, 154–164.

Dickie, E.W., et al., 2018. Personalized intrinsic network topography mapping and functional connectivity deficits in autism spectrum disorder. Biol. Psychiatry 84, 278–286.

Downey, D., et al., 2016. Comparing the actions of lanicemine and ketamine in depression: key role of the anterior cingulate. Eur. Neuropsychopharmacol. 26, 994–1003.

Doyle, O.M., et al., 2013. Quantifying the attenuation of the ketamine pharmacological magnetic resonance imaging response in humans: a validation using antipsychotic and glutamatergic agents. J. Pharmacol. Exp. Ther. 345, 151–160.

Drevets, W.C., Savitz, J., Trimble, M., 2008. The subgenual anterior cingulate cortex in mood disorders. CNS Spectr. 13, 663–681.

Driesen, N.R., et al., 2013. The impact of NMDA receptor blockade on human working memory-related prefrontal function and connectivity. Neuropsychopharmacology 38, 2613–2622.

Duman, R.S., Sanacora, G., Krystal, J.H., 2019. Altered connectivity in depression: GABA and glutamate neurotransmitter deficits and reversal by novel treatments. Neuron 102, 75–90.

Eickhoff, S.B., et al., 2016. Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. Neuroimage 137.

Eklund, A., Nichols, T.E., Knutsson, H., 2016. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. Proc. Natl. Acad. Sci. USA 113, 7900–7905.

Finn, E.S., et al., 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat. Neurosci. 18, 1664–1671.

Friston, K.J., Williams, S., Howard, R., Frackowiak, R.S., Turner, R., 1996. Movement-related effects in fMRI time-series. Magn. Reson. Med. 35, 346–355.

Gärtner, M., et al., 2019. Functional connectivity between prefrontal cortex and subgenual cingulate predicts antidepressant effects of ketamine. Eur. Neuropsychopharmacol. 29, 501–508.

George, J.R., Michaels, T.I., Sevelius, J., Williams, M.T., 2019. The psychedelic renaissance and the limitations of a White-dominant medical framework: A call for indigenous and ethnic minority inclusion. J. Psychedelic Stud. 4, 4–15.

Gerhard, D.M., et al., 2020. GABA interneurons are the cellular trigger for ketamine's rapid antidepressant actions. J. Clin. Invest 130, 1336–1349.

Girn, M., et al., 2023. A complex systems perspective on psychedelic brain action. Trends Cogn. Sci. 27, 433–445.

Girn, M.A., et al., 2022. Serotonergic psychedelic drugs LSD and psilocybin reduce the hierarchical differentiation of unimodal and transmodal cortex. Neuroimage 256, 119220.

Gratton, C., et al., 2020. Removal of high frequency contamination from motion estimates in single-band fMRI saves data without biasing functional connectivity. Neuroimage 217, 116866.

Grinspoon, L., Bakalar, J.B., Others, 1979. Psychedelic drugs reconsidered, vol. 168. Basic Books,, New York.

Head, M.L., Holman, L., Lanfear, R., Kahn, A.T., Jennions, M.D., 2015. The extent and consequences of p-hacking in science. e1002106 PLoS Biol. 13. e1002106.

Höflich, A., et al., 2015. Ketamine-Induced Modulation of the Thalamo-Cortical Network in Healthy Volunteers As a Model for Schizophrenia. Int. J. Neuropsychopharmacol. 18.

Höflich, A., et al., 2017. Ketamine-dependent neuronal activation in healthy volunteers. Brain Struct. Funct. 222, 1533–1542.

Javitt, D.C., et al., 2018. Utility of Imaging-Based Biomarkers for Glutamate-Targeted Drug Development in Psychotic Disorders: A Randomized Clinical Trial. JAMA Psychiatry 75, 11–19.

Joules, R., et al., 2015. Ketamine induces a robust whole-brain connectivity pattern that can be differentially modulated by drugs of different mechanism and clinical profile. Psychopharmacology 232, 4205–4218.

Keller, A.S., et al., 2023. Hierarchical functional system development supports executive function. Trends Cogn. Sci. 27, 160–174.

Kyzar, E.J., Nichols, C.D., Gainetdinov, R.R., Nichols, D.E., Kalueff, A.V., 2017. Psychedelic Drugs in Biomedicine. Trends Pharmacol. Sci. 38.

Lebedev, A.V., et al., 2015. Finding the self by losing the self: Neural correlates of ego-dissolution under psilocybin. Hum. Brain Mapp. 36, 3137–3153.

Li, M., et al., 2020. Default mode network connectivity change corresponds to ketamine's delayed glutamatergic effects. Eur. Arch. Psychiatry Clin. Neurosci. 270, 207–216.

Lord, L.-D., et al., 2019. Dynamical exploration of the repertoire of brain networks at rest is modulated by psilocybin. Neuroimage 199, 127–142.

Luppi, A.I., et al., 2022. Mapping pharmacologically-induced functional reorganisation onto the brain's neurotransmitter landscape. bioRxiv. https://doi.org/10.1101/2022.07.12.499688.

Luscher, B., Feng, M., Jefferson, S.J., 2020. Antidepressant mechanisms of ketamine: Focus on GABAergic inhibition. Adv. Pharm. 89, 43–78.

Ly, C., et al., 2018. Psychedelics promote structural and functional neural plasticity. Cell Rep. 23, 3170–3182.

Lynch, C.J., et al., 2020. Prevalent and sex-biased breathing patterns modify functional connectivity MRI in young adults. Nat. Commun. 11, 5290.

Makowski, C., Lepage, M., Evans, A.C., 2019. Head motion: the dirty little secret of neuroimaging in psychiatry. J. Psychiatry Neurosci. 44, 62–68.

Marek, S., et al., 2022. Reproducible brain-wide association studies require thousands of individuals. Nature 603, 654–660.

Margulies, D.S., et al., 2016. Situating the default-mode network along a principal gradient of macroscale cortical organization. Proc. Natl. Acad. Sci. U. S. A. 113, 12574–12579.

Markello, R.D., et al., 2022. neuromaps: structural and functional interpretation of brain maps. Nat. Methods 19, 1472–1479.

McCulloch, D.E.-W., et al., 2021. Psychedelic resting-state neuroimaging: a review and perspective on balancing replication and novel analyses. PsyArXiv. https://doi.org/10.31234/osf.io/64kyg.

McMillan, R., et al., 2020. Simultaneous EEG/fMRI recorded during ketamine infusion in patients with major depressive disorder. Prog. Neuropsychopharmacol. Biol. Psychiatry 99, 109838.

Michaels, T.I., Purdon, J., Collins, A., Williams, M.T., 2018. Inclusion of people of color in psychedelic-assisted psychotherapy: a review of the literature. BMC Psychiatry 18, 245.

Mitchell, J.M., et al., 2021. MDMA-assisted therapy for severe PTSD: a randomized, double-blind, placebo-controlled phase 3 study. Nat. Med. 27, 1025–1033.

Moher, D., et al., 2015. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst. Rev. 4, 1–9.

Mueller, F., et al., 2017. Acute effects of LSD on amygdala activity during processing of fearful stimuli in healthy subjects. Transl. Psychiatry 7, e1084.

Müller, E.J., et al., 2020. Core and matrix thalamic sub-populations relate to spatio-temporal cortical connectivity gradients. Neuroimage 222, 117224.

Müller, F., Dolder, P.C., Schmidt, A., Liechti, M.E., Borgwardt, S., 2018. Altered network hub connectivity after acute LSD administration. Neuroimage Clin. 18, 694–701.

O'Leary, D.D.M., Chou, S.-J., Sahara, S., 2007. Area patterning of the mammalian cortex. Neuron 56, 252–269.

Pearson, C., Siegel, J., Gold, J.A., 2022. Psilocybin-assisted psychotherapy for depression: Emerging research on a psychedelic compound with a rich history. J. Neurol. Sci. 434, 120096.

Poldrack, R.A., et al., 2008. Guidelines for reporting an fMRI study. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2007.11.048.

Poldrack, R.A., et al., 2017. Scanning the horizon: towards transparent and reproducible neuroimaging research. Nat. Rev. Neurosci. 18, 115–126.

Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59, 2142–2154.

Power, J.D., Plitt, M., Laumann, T.O., Martin, A., 2017. Sources and implications of whole-brain fMRI signals in humans. Neuroimage 146, 609–625.

Preller, K.H., et al., 2018a. Changes in global and thalamic brain connectivity in LSD-induced altered states of consciousness are attributable to the 5-HT2A receptor. Elife 7.
Preller, K.H., et al., 2018b. Role of the 5-HT Receptor in Self- and Other-Initiated Social Interaction in Lysergic Acid Diethylamide-Induced States: A Pharmacological fMRI Study. J. Neurosci. 38, 3603–3611.
Preller, K.H., et al., 2020. Psilocybin induces time-dependent changes in global functional connectivity. Biol. Psychiatry 88, 197–207.
Roiser, J.P., et al., 2016. Minimum statistical standards for submissions to neuroimage. Clin. *Neuroimage Clin.* 12, 1045–1047.
Roseman, L., Leech, R., Feilding, A., Nutt, D.J., Carhart-Harris, R.L., 2014. The effects of psilocybin and MDMA on between-network resting state functional connectivity in healthy volunteers. Front. Hum. Neurosci. 8, 204.
Sadraee, A., Paulus, M., Ekhtiari, H., 2021. fMRI as an outcome measure in clinical trials: A systematic review in clinicaltrials.gov. Brain Behav. 11, e02089.
Satterthwaite, T.D., et al., 2012. Impact of in-scanner head motion on multiple measures of functional connectivity: relevance for studies of neurodevelopment in youth. Neuroimage 60, 623–632.
Satterthwaite, T.D., et al., 2019. Motion artifact in studies of functional connectivity: characteristics and mitigation strategies. Hum. Brain Mapp. 40, 2033–2051.
Scheidegger, M., et al., 2012. Ketamine decreases resting state functional network connectivity in healthy subjects: implications for antidepressant drug action. PLoS One 7, e44799.
Schmidt, A., et al., 2017. Comparative effects of methylphenidate, modafinil, and mdma on response inhibition neural networks in healthy subjects. Int. J. Neuropsychopharmacol. 20, 712–720.
Schmidt, A., et al., 2018a. Acute LSD effects on response inhibition neural networks. Psychol. Med. 48, 1464–1473.
Schmidt, A., et al., 2018b. Acute effects of methylphenidate, modafinil, and MDMA on negative emotion processing. Int. J. Neuropsychopharmacol. 21, 345–354.
Siegel, J.S., et al., 2014. Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. Hum. Brain Mapp. 35, 1981–1996.
Siegel, J.S., et al., 2017. Data quality influences observed links between functional connectivity and behavior. Cereb. Cortex 27, 4492–4502.
Siegel, J.S., Nicol, G.E., 2023. Plasticity markers in the human brain associated with rapid antidepressants. Neuropsychopharmacology 48, 223–224.
Siegel, J.S., Daily, J.E., Perry, D.A., Nicol, G.E., 2023. Psychedelic Drug Legislative Reform and Legalization in the US. JAMA Psychiatry 80, 77–83.
Smith, S.M., Nichols, T.E., 2009. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage 44, 83–98.
Steffens, M., et al., 2016. Effects of ketamine on brain function during smooth pursuit eye movements. Hum. Brain Mapp. 37, 4047–4060.
Sydnor, V.J., et al., 2021. Neurodevelopment of the association cortices: Patterns, mechanisms, and implications for psychopathology. Neuron 109, 2820–2846.
Tagliazucchi, E., et al., 2016. Increased global functional connectivity correlates with LSD-induced ego dissolution. Curr. Biol. 26, 1043–1050.
Thrul, J., Garcia-Romeu, A., 2021. Whitewashing psychedelics: racial equity in the emerging field of psychedelic-assisted mental health research and treatment. Drug.: Educ. Prev. Policy 28, 211–214.
Turesky, T.K., Vanderauwera, J., Gaab, N., 2021. Imaging the rapidly developing brain: current challenges for MRI studies in the first five years of life. Dev. Cogn. Neurosci. 47, 100893.
Van Dijk, K.R.A., Sabuncu, M.R., Buckner, R.L., 2012. The influence of head motion on intrinsic functional connectivity MRI. Neuroimage 59, 431–438.
Vollenweider, F.X., Smallridge, J.W., 2022. Classic psychedelic drugs: update on biological mechanisms. Pharmacopsychiatry 55, 121–138.
Vue, T.Y., et al., 2013. Thalamic control of neocortical area formation in mice. J. Neurosci. 33, 8442–8453.
Wager, T.D., et al., 2013. An fMRI-based neurologic signature of physical pain. New Engl. J. Med. 368, 1388–1397.
Wong, J.J., O'Daly, O., Mehta, M.A., Young, A.H., Stone, J.M., 2016. Ketamine modulates subgenual cingulate connectivity with the memory-related neural circuit-a mechanism of relevance to resistant depression? PeerJ 4, e1710.
Woo, C.-W., Krishnan, A., Wager, T.D., 2014. Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2013.12.058.
Zalesky, A., Fornito, A., Bullmore, E.T., 2010. Network-based statistic: identifying differences in brain networks. Neuroimage 53, 1197–1207.