nature portfolio

Corresponding author(s): Boris Heifets, MD, PhD

Last updated by author(s): 08/29/2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Screening and outcomes data were entered into Stanford REDCap (version 13.4.10), a secure online data capture platform (http://redcap.stanford.edu) developed and operated by Stanford Medicine Research IT team.
Data analysis	RStudio (version 2022.07.1 for MacOS) was used for data analysis. Prism GraphPad (version 9.5.0 for MacOS) was used for figure-making. R code used for data analysis is publicly available at https://osf.io/zdkr8/ (DOI 10.17605/OSF.IO/ZDKR8).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

De-identified participant data, data dictionaries, study protocol, and statistical analysis plan are available at https://osf.io/zdkr8/ (DOI 10.17605/OSF.IO/ZDKR8). All participants have consented to sharing de-identified data with outside entities for scientific research purposes.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Information about birth sex, but not gender, was collected for demographic characterization of our sample. We obtained data on participant birth sex from electronic health records, which rely on patient self report. This data is summarized as a number and percentage for each study arm in the demographics table.
Population characteristics	The mean age of trial participants was 51 years; they were mostly female (70%), white (65%), non-Hispanic (87.5%), employed (62.5%), and never smoked (65%). At screening, both groups had moderate levels of depression (ketamine: mean Montgomery-Asberg Depression Rating Scale score = 27.7, placebo: 30.6) and moderate levels of treatment resistance (ketamine: mean Maudsley Staging Method score = 8.3, placebo: 7.5).
Recruitment	Adults undergoing elective non-cardiac, non-intracranial surgery were recruited from preoperative clinics at Stanford University Medical Center. The 8-item Patient Health Questionnaire (PHQ-8) was distributed to patients through a perioperative mental health screening service. To be eligible for the study, patients must score ≥ 10 on the PHQ-8, corresponding with at least moderate depression. Research staff screened electronic health records (EHR) of patients scheduled for surgery who scored ≥ 10 points on the PHQ-8; those without exclusion criteria documented in the EHR were introduced to the study and consented for an additional screening visit. Surgical clinics also referred patients with symptomatic depression who expressed interest in the trial. These patients were contacted by research staff for a telephone pre-screen, and qualifying patients were consented for an additional screening visit. At this visit, research staff collected information on demographics, medical, and psychiatric history, including level of antidepressant treatment resistance assessed by the Maudsley Staging Method (MSM). Inclusion and exclusion criteria were assessed via a hybrid approach of corroborating EHR data with patient self-report. Self-selection bias may have favored patients with less severe depression, as these patients are more likely to engage with research staff and complete all screening procedures. However, our eligibility criteria set minimum depression severity scores to reduce the impact of this type of bias.
Ethics oversight	This was an investigator-initiated, university-sponsored trial. The trial protocol was approved by the institutional review board at Stanford University and all participants gave written informed consent.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Our sample size estimation was derived from an a priori power analysis for the primary outcome. In a randomized controlled trial of ketamine versus active placebo performed by Phillips et al., participants had a mean decrease of 10.9 points (standard deviation [SD] 8.9) in MADRS total score relative to pre-infusion scores compared with a mean decrease of 2.8 points (SD 3.6) with midazolam29. For reference, the minimum clinically important difference on the MADRS is estimated to range from 3 to 9 points30. Using these results, we computed an estimated total sample size of 38 participants at a two-sided alpha level of 0.05 and 80% power to detect this difference if using parametric testing. An additional 2 participants were added to account for potential attrition, for a total of 40 participants. Interim analyses were not performed.
Data exclusions	No data was excluded for primary and secondary and exploratory analyses.
Replication	At every screening and follow-up time point, all outcomes were independently measured once. No replication experiments were conducted.

Randomization	Twenty participants were randomly allocated to a single dose of intravenous ketamine (0.5 mg/kg diluted into 40 ml of normal saline, infused over 40 minutes using a programmable pump). Another twenty participants were randomly allocated to 40 ml of normal saline infused similarly over 40 minutes. Pharmacy staff randomized participants using computerized block randomization with 5 blocks of 8.
Blinding	The participant, investigators, and direct care providers (e.g., anesthesiologists) were masked to treatment assignment. Prior to participant unmasking at the end of the follow up period, a masking assessment of participants was performed.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	NCT03861988
Study protocol	The full trial protocol will be submitted along with the manuscript and SAP.
Data collection	Data was collected by 1) review of electronic health records, 2) in-person interviews on Stanford University campus and Stanford Hospital, 3) telephone and video-based interviews. Participant recruitment occurred between February 2020 and August 2022. Data collection occurred between February 2020 and September 2022.
Outcomes	The primary outcome was the MADRS score measured 1, 2, and 3 days post-infusion, as previous studies have found the greatest antidepressant effect occurs within the first 72 hours of a single ketamine infusion. The MADRS is a clinical rating scale used widely in antidepressant trials; it consists of 10 items which measure the severity of depression in individuals, with a total score ranging from 0 to 60, and higher scores indicating more severe depression15. Trained clinical research personnel administered the MADRS. The secondary outcome was clinical response, defined as ≥50% reduction in MADRS scores from screening baseline. Remission, defined as MADRS score ≤12 in our study, was treated as an exploratory outcome.