nature portfolio

Corresponding author(s): Erping Long
Qingyu Chen

Last updated by author(s): 2024/06/09

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	During the data collection, the software iFLYTEK (v3.0.1734) was used for the audio-to-text transformation.
Data analysis	The model was developed with GPT3.5, OpenAI v0.28.1 (https://github.com/openai/openai-python), Ragas v0.0.18 (https://github.com/explodinggradients/ragas), Langchain v0.0.333 (https://github.com/langchain-ai/langchain). Data analysis was done with Microsoft Excel. The source code could be accessed via the following link: https://github.com/ZigengHuang/SSPEC.

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

The data supporting the findings of this trial are available within the manuscript and its Supplementary Information files. All requests for further data sharing will be reviewed by the Ethics Review Committee of Southern University of Science and Technology Yantian Hospital, Shenzhen, China, and Renmin Hospital of Wuhan

University, Wuhan, China, to verify whether the request is subject to any intellectual property or confidentiality obligations. Requests for access to de-identified participant-level data from this trial can be submitted via email to Dr. Erping Long (erping.long@ibms.pumc.edu.cn) with detailed proposals for approval and will be responded to within 60days. Each participant's rights and privacy are key subjects taken into consideration when sharing information. A signed data access agreement with the collaborator is required before accessing shared data. The raw conversation data are not publicly available due to restrictions of privacy.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Gender information were self-reported by the participants. Both genders of participants were recruited. The clinical trial enrolled 542 (50.2%) male participants in the nurse group and 526 (48.5%) male participants in the nurse-SSPEC group.
Reporting on race, ethnicity, or other socially relevant groupings	No groupings of race, ethnicity, or socially relevant factors were involved in this study.
Population characteristics	Population characteristics are provided in Table 1.
Recruitment	Potential participants were recruited by the research team if they meet the inclusion criteria. The participants have signed the informed consent and understood that this is an exploratory experiment, which should not be interpreted as direct guidance for clinical interventions at this stage. Eligible participants were recruited at the discretion of the investigator. No selection bias was considered to be present in recruitment process. Recruitment was limited to locations or regions where the clinical trial was being conducted.
Ethics oversight	The study protocol and all amendments were approved by the ethics review committee of the Chinese Academy of Medical Sciences and other participating centers. We obtained informed consent from all participants (both patients and nurse in this study and the tenets of the Declaration of Helsinki were followed throughout this study. This study implemented stringent data protection measures, ensuring that all data was anonymized and encrypted to protect privacy.

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	Sample size were estimated based on satisfaction difference between human and SSPEC in internal validation as the reference rate. We calculated that a sample size of 1,696 participants (assuming a 1:1 allocation ratio) would be required to achieve 80% power at a 0.05 level of significance (β = 0.2) based on the reference satisfaction rate in the internal validation.
Data exclusions	In clinical trial, the participants (n=12 in nurse-SSPEC group and n=9 in nurse group) who requested the withdrawal or discontinuation were excluded from the final analysis.
Replication	Not available in this study as no lab experiments is invovled.
Randomization	In the clinical trial, each participant was randomly assigned to either nurse group or the nurse-SSPEC group.
Blinding	In the internal validation and the clinical trial, the evaluators were blind to the source of response (nurse or SSPEC).

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
☒	☐ Plants

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	The clinical trial was prospectively registered in Chinese Clinical Trial Registry (identifier: ChiCTR2300077245).
Study protocol	The study protocol and statistical analysis plan were provided with the submission
Data collection	Between November 8, 2023 and November 24, 2023, a total of 2185 individuals were enrolled in the Southern University of Science and Technology Yantian Hospital in Shenzhen city. The audio data between receptionist nurses and patients were collected from 6 reception sites: general service at the hospital entrance, emergency at entrance of the emergency room, as well as internal medicine, surgery, gynecology, and pediatrics located near their respective departments.
Outcomes	Primary outcome: Patient's satisfaction scale to the responses from either clinician-SSPEC or clinician model. Secondary outcomes: 1) Rate of negative emotional expressions and recurrent query-and-response patterns. 2) Response quality in terms of factuality, integrity, safety, empathy, and readability. 3) Clinician's satisfaction rate to collaborate with SSPEC.

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*