

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  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The exact sample size ( <i>n</i> ) for each experimental group/condition, given as a discrete number and unit of measurement   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A description of all covariates tested   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> 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) |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i> ) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted<br><i>Give P values as exact values whenever suitable.</i>                     |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i> ), 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	Datasets 1-3, 8 were collected using Eprime 2.0 (Psychology Software Tools Inc., Pittsburgh, USA). Datasets 4-7 were collected using PsychoPy (version: py3.8, Peirce et al., 2019). The timing of all experiments was synchronized with the MRI system by waiting for the presentation of a blank screen upon pressing the 's/S' key before the start of the experimental procedure.
Data analysis	All image preprocessing was performed using deepbet (version 0.0.2), Spinal Cord Toolbox (SCT; version 6.1; <a href="https://spinalcordtoolbox.com/">https://spinalcordtoolbox.com/</a> ), FMRIB Software Library (FSL; version 6.0.5; <a href="https://fsl.fmrib.ox.ac.uk/">https://fsl.fmrib.ox.ac.uk/</a> ), and custom scripts written in Python (version 3.9.18). Subsequent time-series analyses were conducted using NumPy (version 1.26.4; <a href="https://numpy.org/">https://numpy.org/</a> ), SciPy (version 1.13.1; <a href="https://www.scipy.org/">https://www.scipy.org/</a> ), antspyx (version 0.4.2; <a href="https://github.com/ANTsX/antspyx">https://github.com/ANTsX/antspyx</a> ), nibabel (version 5.2.1; <a href="https://nipy.org/nibabel/">https://nipy.org/nibabel/</a> ), nilearn (version 0.10.3; <a href="https://nilearn.org/">https://nilearn.org/</a> ), and scikit-learn (version 1.3.0; <a href="https://scikit-learn.org/">https://scikit-learn.org/</a> ). All tools were implemented in a Python 3.9.18 environment. Custom code developed for this study is openly available on Zenodo at: <a href="https://doi.org/10.5281/zenodo.15420790">https://doi.org/10.5281/zenodo.15420790</a> .

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 corticospinal fMRI time-series data generated in this study have been deposited in Zenodo under the identifier <https://doi.org/10.5281/zenodo.15469454>. All raw data underlying the figures and extended data are provided in the accompanying Source Data file, which reproduces all plots in the manuscript. The DICOM imaging data are available under controlled access due to participant privacy and applicable data-protection regulations. Qualified researchers may request de-identified DICOMs for research, non-commercial use by (i) providing evidence of institutional ethics approval covering the proposed use, and (ii) signing a Data Use Agreement that prohibits re-identification, redistribution, and use beyond the approved purpose. Access requests should be submitted to Prof. Yazhuo Kong ([kongyz@psych.ac.cn](mailto:kongyz@psych.ac.cn)); requests will be reviewed and responded to within 15 working days.

## 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

The study included participants of female and male, with sex classification based on self-reported information. No sex-specific analyses were conducted, as the primary objective focused on identifying population-level neural signatures of pain sensitivity.

### Reporting on race, ethnicity, or other socially relevant groupings

All participants were of East Asian descent. The study protocol did not include specific ethnic background data collection, and all analyses were performed on neuroimaging and behavioral datasets without racial/ethnic stratification.

### Population characteristics

Dataset 1: 54 healthy participants (31 females; age =  $22.25 \pm 3.13$  years)  
 Dataset 2: 48 healthy participants (20 females; age =  $22.06 \pm 2.99$  years)  
 Dataset 3: 36 healthy participants (16 females; age =  $21.47 \pm 3.48$  years)  
 Dataset 4: 32 healthy participants (20 females; age =  $23.42 \pm 2.96$  years)  
 Dataset 5: 22 type 2 diabetic patients (9 females; age =  $59.68 \pm 7.95$  years)  
 Dataset 6: 24 ZAP patients (7 females; age =  $55.79 \pm 11.82$  years)  
 Dataset 7: 58 healthy participants (30 females; age =  $23.43 \pm 2.62$  years; )  
 Dataset 8: 462 healthy participants (282 females; age =  $21.51 \pm 4.64$  years)

### Recruitment

Healthy volunteers were recruited through local advertisements. Exclusion criteria included: (1) history of chronic pain, mental illness, or regular analgesic use; (2) current use of pain-relief medications; and (3) contraindications to MRI procedures.  
 Patient participants were recruited by clinical specialists from hospital outpatient departments. Exclusion criteria comprised: (1) a history of cerebral tumors, neurosurgical procedures, or cranial trauma; (2) presence of acute complications; (3) severe hypertension; (4) a significant history of neurological, psychiatric, or cerebrovascular disorders; (5) history of alcohol or substance abuse; and (6) with mental disorders or contraindications to MRI.

### Ethics oversight

The institutional review boards of Institute of Psychology, Chinese Academic of Science (No. H19023, H20055, H24051), Shandong Provincial Hospital (No. 2024-398), Beijing Xuanwu hospital (No. 2024-005-001) approved the study protocols.

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](https://www.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

No formal power calculations or statistical tests were performed to predetermine the sample size. However, the chosen sample size aligns with established standards in the field and is consistent with prior methodological recommendations (e.g., Lee et al., 2021; Kohoutová et al., 2020). To ensure the robustness and generalizability of the results, analyses were extended to include seven independent replication datasets.

### Data exclusions

In Dataset 1, 9 participants were excluded due to missing thermal pain threshold data, and 2 participants were excluded for incomplete spinal imaging data. In Dataset 3, 3 participants lacked spinal cord imaging data, and 1 participant was excluded for excessive head movement (>3mm) during scanning. In Dataset 4, 1 participant was excluded for excessive head movement (>3mm). In Dataset 5, 10 participants with type 2 diabetes were excluded due to the absence of chronic pain. In Dataset 7, 2 healthy participant was removed for the lack of spinal cord

signal. In Dataset 8, 3 participants were excluded for excessive head motion, 7 participants were removed due to missing cold pain threshold behavioral data, and 12 participants were excluded as data outliers (beyond 3 standard deviations).

Replication	Model validation: The trained model was evaluated using leave-one-out cross-validation on the primary dataset and further tested for generalizability through independent validation, pain-specific validation, and comparative model benchmarking across seven additional datasets. Results demonstrated robust replication of effects across individuals and studies, confirming the model's consistency and validity.
Randomization	Not applicable.
Blinding	For safety reasons, blinding of the TMS operator was not feasible in Dataset 7; however, participants and all personnel involved in outcome assessment and data analysis remained blinded to group allocation throughout data collection and data pre-processing. To preserve participant blinding, the sham coil was indistinguishable from the active coil in external appearance and reproduced the characteristic acoustic output of stimulation, while not delivering effective cortical stimulation.

## 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

### Methods

n/a	Involved in the study	n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies	<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines	<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology	<input type="checkbox"/>	<input checked="" type="checkbox"/> MRI-based neuroimaging
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants		

### 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, mosaicism, off-target gene editing) were examined.

## Magnetic resonance imaging

### Experimental design

Design type	Corticospinal fMRI
Design specifications	Dataset 1-3, 5-7: Resting-state corticospinal fMRI Dataset 4: a heat pain task-based and a cold sensation task-based corticospinal fMRI
Behavioral performance measures	Dataset 1-3: Thermal pain sensitivity thresholds Dataset 4: Heat/cold sensation intensity ratings (VAS 0-10) Dataset 5: Chronic pain severity (McGill Pain Questionnaire, MPQ) Dataset 6: Acute pain intensity (NRS 0-10) Dataset 7: Offline heat-pain task ratings (VAS 0-10) Dataset 8: Cold pain sensitivity thresholds

### Acquisition

Imaging type(s)	Functional
Field strength	3T
Sequence & imaging parameters	We utilized a modified corticospinal resting-state and task-based fMRI protocol based on the Siemens SMS-EPI

sequence to simultaneously capture the brain and cervical cord's BOLD response. The field of view (FOV) was expanded to 192×192mm<sup>2</sup>, covering the entire brain and cervical spinal cord (C1-C7) with 70 slices, each 4 mm thick. A 3D shimming volume focused on the brainstem and cervical cord, with the phase coding direction set from posterior to anterior (P→A), flip angle at 80°, FOV phase at 100%, and in-plane resolution at 1.5 × 1.5 mm<sup>2</sup>. To enhance efficiency, a simultaneous multiband GRAPPA parallel imaging technique with an interleaved multi-slice mode and an acceleration factor of 2 was applied, achieving a repetition time (TR) of 2680 ms and an echo time (TE) of 27 ms. In total, 200 volumes were scanned for Dataset 1~3, 5, and 235 volumes were scanned for Dataset 4, 6. During the resting-state fMRI scanning, participants were instructed to relax and remain still with their eyes open, not to fall asleep, and not to think about anything.

To correct distortion around the brain and cervical spine caused by the extended FOV and magnetic field inhomogeneities, two opposing phase encoding B0 images were employed (for details, see data preprocessing; for correction effect, see Wei et al., 2024). These images, with TR: 7500 ms and TE: 87 ms, used a 2D SP-EPI sequence, maintaining the protocol's FOV, resolution, echo spacing, acceleration, and shimming.

Brain MRI data (Dataset 8) were collected using a 3.0-Tesla MRI system (Discovery MR 750; General Electric Healthcare, Milwaukee, WI, USA) with 8-channel head coil and the restraining foam pads to minimize head motion and scanner noise. High-resolution T1-weighted structural brain images were acquired using a gradient echo (3D SPGR) sequence with the parameters included sagittal slice orientation, TR: 2000 ms, TE: 2 ms, TI: 450 ms, FOV: 176 × 256 × 256 mm<sup>3</sup>, FOV phase: 100%, and flip angle: 8°. Resting-state functional images were collected over ten minutes using an echo-planar imaging sequence with parameters included TR: 2000 ms, TE: 29 ms, flip angle: 90°, field of view: 192 × 192 mm<sup>2</sup>, in-plane resolution: 3 × 3 mm<sup>2</sup>, slice thickness: 3 mm.

Area of acquisition

Corticospinal

Diffusion MRI

☐ Used

☒ Not used

## Preprocessing

Preprocessing software

FSL, AFNI, antspxy, SCT

Normalization

For brain data, two-step registration was performed using ANTs: functional to structural, then structural to standard space, utilizing ANTs' 'SyNRA' and 'SyBoldAff' algorithms. For spinal cord data, SCT was used for normalization.

Normalization template

Brain: MNI152 (2mm); Spinal cord: PAM50

Noise and artifact removal

For noise and artifact removal, physiological noise from cardiac and respiratory cycles was mitigated via FSL PNM. Susceptibility-induced distortions in fMRI images were corrected using FSL's topup with opposing-phase-encoding (AP/PA) B0 field maps. Head motion artifacts were addressed through rigid-body motion correction (MCFLIRT), and motion-related noise components were further suppressed using FSL-AROMA to automatically regress out motion-corrupted ICA components. Tissuespecific confounds were removed by regressing out cerebrospinal fluid (CSF) and white matter (WM) signals derived from FSL's FAST segmentation. For spinal cord data, additional CS pulsation noise was modeled using the top 10% variance voxels in CSF masks, alongside motion parameters from X-Y translations.

Volume censoring

Data was excluded for excessive head movement (>3mm).

## Statistical modeling & inference

Model type and settings

Individual-level statistical analyses employed atlas-based partial correlation, while group-level inferences employed multivariate pattern decoding to assess cross-participant consistency and distributed neural representations.

Effect(s) tested

Pearson correlation coefficients were computed across participants to quantify the linear relationship between CSps responses and behavioral measures of pain sensitivity. To evaluate the statistical significance of these predictive relationships, a nonparametric bootstrapping procedure was implemented to generate null distributions of correlation coefficients under the hypothesis of no association. The empirical p-value was calculated as the proportion of bootstrapped iterations where the resampled correlation coefficient exceeded the observed value, testing against the null hypothesis.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☒ Both

Anatomical location(s)

122 regions of interest based on the Multiresolution Intrinsic Segmentation Template brain atlas, PAG, RVM, 14 spinal cord horns, and whole grey matter

Statistic type for inference

functional connectivity based on MIST atlas

(See [Eklund et al. 2016](#))

Correction

false discovery rate

## Models &amp; analysis

n/a	Involvement in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Functional and/or effective connectivity
<input checked="" type="checkbox"/>	<input type="checkbox"/> Graph analysis
<input type="checkbox"/>	<input checked="" type="checkbox"/> Multivariate modeling or predictive analysis

## Functional and/or effective connectivity

For resting-state fMRI data, functional connectivity was assessed using partial correlation analysis to estimate neural interactions between regions while controlling for shared signals.

For task-based fMRI data, dynamic functional connectivity was quantified via Dynamic Conditional Correlation (DCC), a time-varying connectivity metric based on generalized autoregressive conditional heteroscedastic (GARCH) and exponentially weighted moving average (EWMA) models ([https://github.com/canlab/Lindquist\\_Dynamic\\_Correlation](https://github.com/canlab/Lindquist_Dynamic_Correlation)).

## Multivariate modeling and predictive analysis

We used the lower triangular matrix of the partial correlation functional connectivity matrices from Dataset 1 as feature values for predictive modeling in machine learning, aiming to predict individual heat pain threshold. Using Python's scikit-learn library (version 1.3.0, <https://github.com/scikit-learn/scikit-learn>), we constructed a machine learning pipeline as previous study<sup>7</sup>. Specifically, the pipeline included robust feature scaling (i.e., removing the median and scaling the data according to the quantile range), K best feature selection (i.e., selecting the number of top features based on their importance or relevance to the pain threshold), and application of the Elastic Net regression model—a linear model combining L1 and L2 regularization. The choice of Elastic Net regression model was based on previous studies<sup>7</sup>, primarily for its ability to adjust sparsity (i.e., the L1 and L2 regularization) as a hyperparameter, avoiding any a priori assumptions about the underlying data sparsity. The free hyperparameters in the machine learning pipeline included the number of pre-selected features (K), the L1/L2 regularization ratio (l1\_ratio), and the regularization weight ( $\alpha$ ). Hyperparameter optimization was achieved through leave-one-out cross-validation, where in each internal validation loop, one participant was held out as validation data and the rest as training data. This cross-validation integrated the complete machine learning process to avoid dependencies between training and test samples. Ultimately, grid search results identified the optimal hyperparameter combination as K = 2000, l1\_ratio = 0.999, and  $\alpha$  = 0.1. The finalized model underwent rigorous validation across multiple datasets, including generalizability testing, specificity testing, and clinical application.