nature neuroscience

Supplementary information

https://doi.org/10.1038/s41593-026-02221-3

Personalized brain decoding of spontaneous pain in individuals with chronic pain

In the format provided by the authors and unedited

SUPPLEMENTARY APPENDIX

PERSONALIZED BRAIN DECODING OF SPONTANEOUS PAIN IN INDIVIDUALS WITH CHRONIC PAIN

Jae-Joong Lee[1], Seongwoo Jo[2], Sungkun Cho[2], Choong-Wan Woo[1,3,4,5*]

[1]Center for Neuroscience Imaging Research, Institute for Basic Science, Suwon, South Korea
[2]Department of Psychology, Chungnam National University, Daejeon, South Korea
[3]Department of Biomedical Engineering, Sungkyunkwan University, Suwon, South Korea
[4]Department of Intelligent Precision Healthcare Convergence, Sungkyunkwan University, Suwon, South Korea
[5]Department of Brain Science and Engineering, Sungkyunkwan University, Suwon, South Korea

[*]Please address correspondence to: waniwoo@skku.edu

This PDF file includes:

SUPPLEMENTARY METHODS

DATA ACQUISITION AND PREPROCESSING

Data acquisition

Imaging was performed using a 3T Siemens Prisma scanner at Sungkyunkwan University. For each session, whole-brain fMRI images were acquired using a gradient-echo EPI sequence with TR = 460 ms, TE = 26 ms, flip angle = 90 degrees, multiband acceleration factor = 8, field of view = 216 mm, 80 × 80 × 56 matrix, 2.7 × 2.7 × 2.7 mm voxels. We also acquired the two spin-echo EPI scans, one with the same phase encoding direction as the fMRI images and the other with the reversed phase encoding direction. For the session 1, 11, and 21, T1-weighted structural images were acquired with TR = 2400 ms, TE = 2.34 ms, TI = 1150 ms, flip angle = 8 degrees, 224 × 320 × 320 matrix, 0.7 × 0.7 × 0.7 mm voxels. For the session 2, 12, and 22, T2-weighted structural images were acquired with TR = 3100 ms, TE = 566 ms, 224 × 320 × 320 matrix, 0.7 × 0.7 × 0.7 mm voxels. For the sessions 3-10, 13-20, and 23-30, diffusion weighted images (DWIs) were acquired with TR = 4100 ms, TE = 89 ms, multiband acceleration factor = 3, field of view = 240 mm, 140 × 140 × 81 matrix, 1.7 × 1.7 × 1.7 mm voxels, 103 volumes (7, 6, 15, 15, 60 images for b-values = 0, 500, 1000, 2000, 3000, respectively). We also acquired the two spin-echo EPI scans, one with the same phase encoding direction as the DWI images and the other with the reversed phase encoding direction.

Preprocessing

Structural and functional MRI data were preprocessed using our in-house preprocessing pipeline as follows. For structural data, we used Freesurfer ('recon-all') to generate cortical surface maps based on T1- and T2-weighted structural images. Then we used ciftify ('ciftify-recon-all') to convert the outputs of Freesurfer, including the cortical surfaces and subcortical volumes, into CIFTI format[1]. During this process, non-linear transformation parameters projecting native T1 space to MNI 2 × 2 × 2 mm^3 template space were also calculated using FSL ('flirt' and 'fnirt'). For functional data, we first removed the initial 22 fMRI image volumes (10 seconds) of each run to allow for image intensity stabilization. Then, the images were motion-corrected using AFNI ('3dvolreg') and distortion-corrected using FSL ('topup' and 'applytopup'). These images were aligned to the first run of the first session using AFNI ('3dvolreg'), co-registered to T1-weighted images using FSL ('flirt' and 'bbregister'), and projected to MNI 2 × 2 × 2 mm^3 template using the non-linear transformation parameters calculated during ciftify. These transformation processes were combined into a single interpolation using FSL ('applywarp').

After the standard preprocessing of functional data, we conducted motion censoring, denoising, and temporal filtering using AFNI ('3dTproject'). First, the motion-contaminated volumes with framewise displacement (FD)[2] > 0.2 mm were removed and replaced through linear interpolation over time. To prevent overly aggressive motion censoring due to respiratory artifacts, head motion parameters were low pass-filtered (< 0.1 Hz) prior to FD calculation[3]. Subsequently, denoising and temporal filtering were carried out in a single nuisance regression step. Regressors for denoising include a linear trend, 6 head motion parameters derived from motion correction[4], 5 principal components of white matter (WM) signals and 5 principal components of cerebrospinal fluid (CSF) signals[5]. For temporal filtering, we applied band-pass filter (0.005 Hz – 0.1 Hz) for the resting condition and high-pass filter (> 0.005 Hz) for the other conditions.

Following the nuisance regression, we used ciftify ('ciftify_subject_fmri') to project the fMRI images to the cortical surface and convert them into CIFTI format. Lastly, the motion-contaminated volumes were excluded from further analyses.

EXPERIMENTAL PROCEDURES

Overview

We conducted a longitudinal, multiple sessions of fMRI experiment involving repeated fMRI scanning. For each fMRI session, we conducted four experimental conditions: (1) resting, (2) spontaneous pain rating, (3) speaking, and (4) listening. The four conditions were conducted consecutively.

Resting condition

The resting condition consisted of a 10-minute run during which participants were lying in the scanner with their eyes fixated on a cross on the screen. This condition was used for deriving individual-specific brain parcels (see Section S3 for details).

Spontaneous pain rating condition

The spontaneous pain rating condition consisted of three 10-minute runs. In this condition, participants continuously reported their moment-by-moment spontaneous pain intensity ratings using a trackball mouse to move a bar on the screen. Participants were instructed to provide these ratings continuously and had the opportunity to practice the rating procedure before the start of this condition. To induce fluctuations in spontaneous pain in a naturalistic yet tolerable manner, participants performed an individualized physical maneuver while lying on the

bed before the start of the second run (see the "Individualized pain exacerbation maneuvers" section below). This condition was used for training and testing personalized decoding models.

Speaking condition

The speaking condition consisted of two 5-minute runs. In this condition, participants freely spoke about one pre-determined question per run while their speech was recorded. These questions were designed to elicit personal stories regarding pain, self, functional disability, social relationship, emotion, and their daily life. In sessions 10 and 20-30, the same two questions were repeated to assess longitudinal changes of semantic features in the speech. This condition was not used for the current study.

Listening condition

The listening condition consisted of two 5-minute runs. In this condition, participants listened to recordings of their speech from the speaking condition while keeping their eyes fixated on a cross on the screen. This condition was not used for the current study.

Individualized pain exacerbation maneuvers

To induce variations in spontaneous pain intensity, we first tested three physical maneuvers in the behavioral experiment room and selected one maneuver that effectively increased spontaneous pain for the fMRI experiment. We used this personalized approach, motivated by a previous study[6]. The selection procedure was as follows.

First, participants provided baseline pain intensity ratings using a visual analog scale (VAS). Subsequently, we implemented three distinct physical maneuvers, each designed to temporarily exacerbate participants' pain while they were lying on the bed. These maneuvers included muscle tightening, pressure application, and joint movement.

During the muscle tightening procedure, participants were instructed to simultaneously tighten the muscles in their forearm, calf, and thigh for as long as possible. This procedure was repeated until the participants reported an increase in pain. In the pressure application procedure, participants reported their most painful limb, and we applied an elastic band on the limb to exert sustained pressure until they reported an increase in pain. In the joint movement procedure, participants were instructed to move their joints by gripping and releasing their hand, rotating their wrist and ankle, or raising their arm and leg until they reported an increase in pain.

Following each procedure, participants provided pain intensity ratings every minute for five minutes. If the pain intensity increased by at least 2 out of 10 in at least one of these five ratings—and they attributed their increased pain to the physical maneuver—we chose that maneuver for the fMRI experiment. If not, we proceeded to the next maneuver in the order of

muscle tightening, pressure application, and joint movement. If no maneuver successfully increased pain, we discussed with the participants to identify an alternative physical maneuver and tested the maneuver until the participants reported an increase in pain.

Internet-delivered psychological intervention

After the 20 sessions of fMRI scanning, participants received an acceptance and commitment therapy (ACT) program for pain management. The ACT program consisted of eight sessions and was delivered online by a licensed therapist. Each session was scheduled between fMRI scanning sessions (e.g., intervention session 1 was delivered between scanning sessions 20 and 21). The effects of ACT were not analyzed in the current study.

INDIVIDUAL-SPECIFIC BRAIN PARCELLATION

We generated individual-specific brain parcellations as in the previous study[7], using publicly available code (https://github.com/MidnightScanClub/MSCcodebase). A brief description of this method is as follows. First, we concatenated all fMRI images from the resting condition and calculated a correlation matrix between all points (i.e., vertices) in the cortex. We then calculated the correlation of this correlation matrix, which represents the similarity of the correlation profile between vertices. Spatial gradients of this similarity matrix were computed, and a watershed algorithm was used to identify edges in the spatial gradients. The edge maps were averaged across all vertices to compute the boundary map of the vertex-wise correlation patterns. Next, we applied the watershed algorithm starting from all local minima to generate brain parcels. If two neighboring parcels had fewer edges than 50th percentile of edge counts across all vertices, they were considered similar and thus merged. The number of resulting parcels were 573, 566, and 572 for Participants 1, 2, and 3, respectively.

To mitigate potential confounds from visual processing of the rating bar, we removed parcels assigned to the medial visual and lateral visual networks (see "Personalized brain network mapping" section), as well as parcels whose vertices overlapped with the cuneus, fusiform, lateral occipital, lingual, or pericalcarine areas, from further analyses. The areal assignments were obtained during the Freesurfer 'recon-all' process based on the Desikan-Killiany atlas[8]. The number of resulting parcels after the removal of visual networks and occipital areas was 448, 445, and 452 for Participant 1, 2, and 3, respectively.

We combined the resulting cortical parcellation with existing subcortical parcellations, including 54 subcortical regions from the Tian atlas[9], 10 cerebellar regions from the multi-domain task battery atlas[10], and the periaqueductal gray and brainstem regions used in previous

studies[11,12]. The final number of whole-brain parcels was 514, 511, and 518 for Participants 1, 2, and 3, respectively.

PERSONALIZED BRAIN NETWORK MAPPING

To identify personalized brain network structures, we conducted Infomap-based community detection as described in a previous study[7], using publicly available code (https://github.com/MidnightScanClub/MSCcodebase). First, we concatenated all fMRI images from the resting condition and calculated a correlation matrix between all vertices/voxels (for vertex- and voxel-wise network structure) or between parcels (for parcel-wise network structure). Correlations between vertices/voxels or parcels whose centroids are within 30 mm of each other were set to zero. Then, we thresholded the correlation matrix across a set of edge densities ranging from 0.3% to 5% and applied Infomap algorithm[13] to each thresholded matrix.

We assigned the network labels to the identified communities based on their similarities to known group-average networks (i.e., default mode network [DMN], anterior medial temporal network [AMTL], contextual association network [CAN], medial visual network [MVIS], lateral visual network [LVIS], parietal memory network [PMN], frontoparietal network [FPN], dorsal attention network [DAN], premotor network [PMOT], language network [LANG], action mode network [AMN], face somatomotor network [FSMN], hand somatomotor network [HSMN], leg somatomotor network [LSMN], and auditory network [AUD]), and a consensus of network assignments was obtained by collapsing across all thresholds.

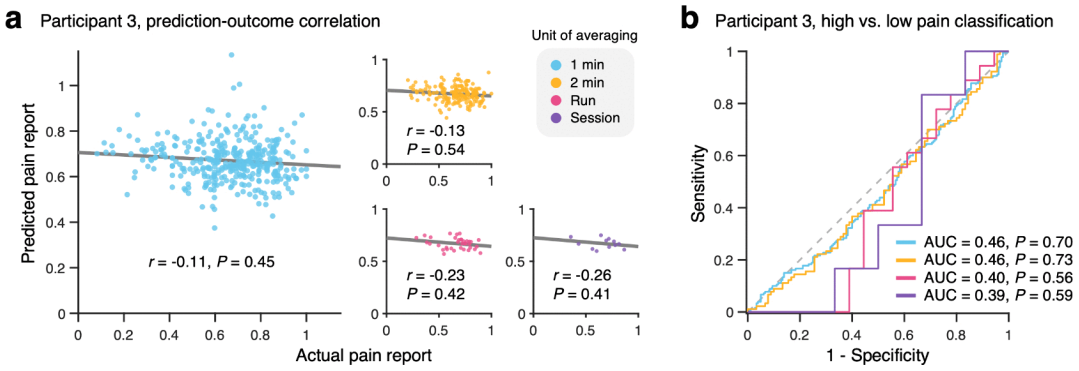
To identify somato-cognitive action network (SCAN)[14], we conducted node-specific thresholding across a set of edge densities ranging from 0.01% to 5%[15]. We manually grouped all inter-effector subnetworks at a density level of 0.1%, which has been reported as an optimal threshold[15]. When inter-effector subnetworks were not evident at a density level of 0.1%, we explored sparser densities (0.05%, 0.02%, and 0.01%) and manually grouped all inter-effector subnetworks as SCAN.

IMPLEMENTATION OF MACHINE LEARNING ALGORITHM

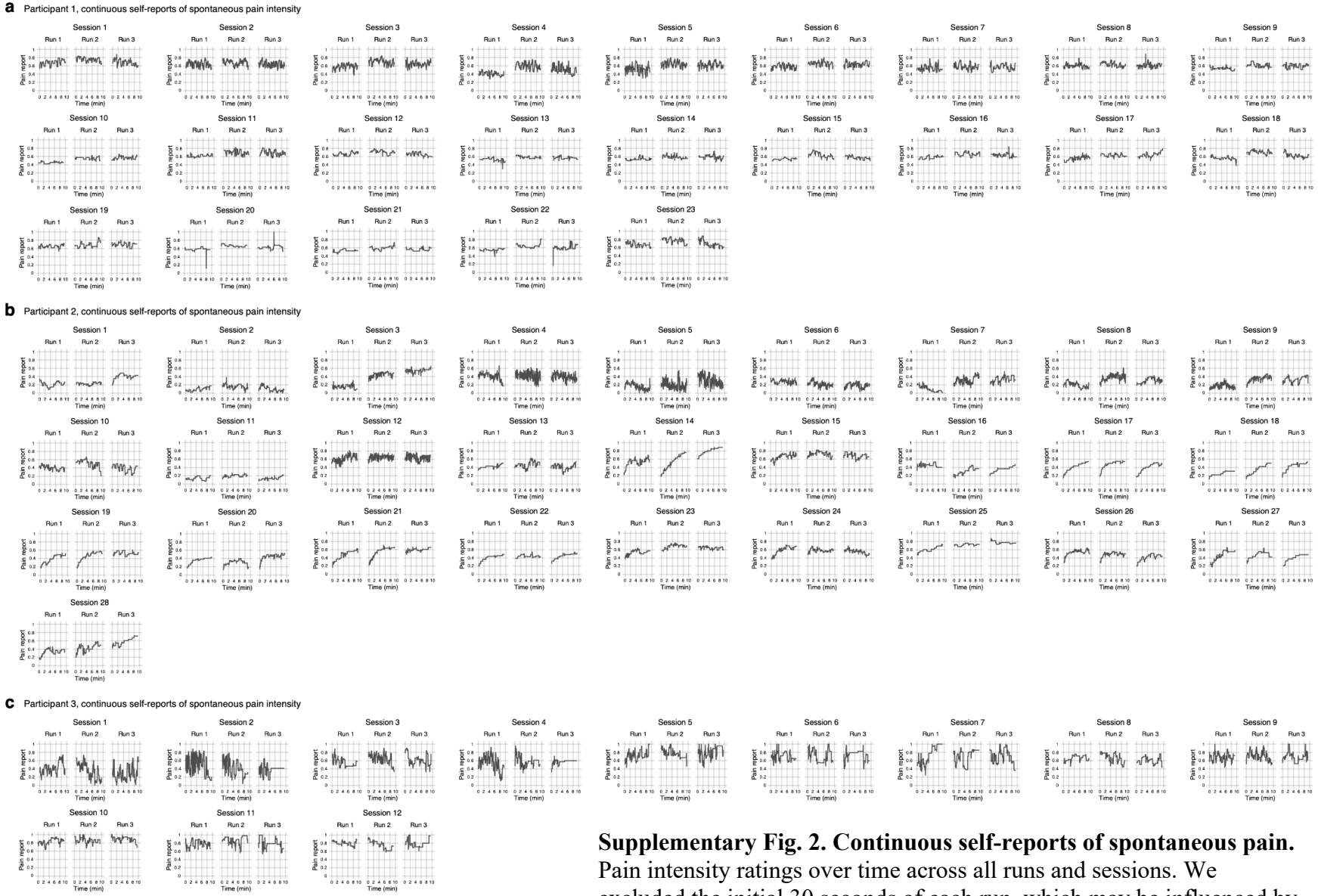
For developing personalized decoding models, we used the LASSO-PCR (least absolute shrinkage and selection operator-regularized principal components regression) algorithm as follows. First, we vectorized and concatenated all binned edge timeseries data from the training data sessions and reduced their dimensionality using Principal Component Analysis (PCA). We then conducted multiple linear regression of pain ratings on the PCA-projected edge timeseries

data with LASSO regularization, which retains only the features important for prediction and shrinks the other features to zero. Then, we selected the principal components (PCs) that survived LASSO regularization and conducted multiple linear regression again using only the selected PCs.

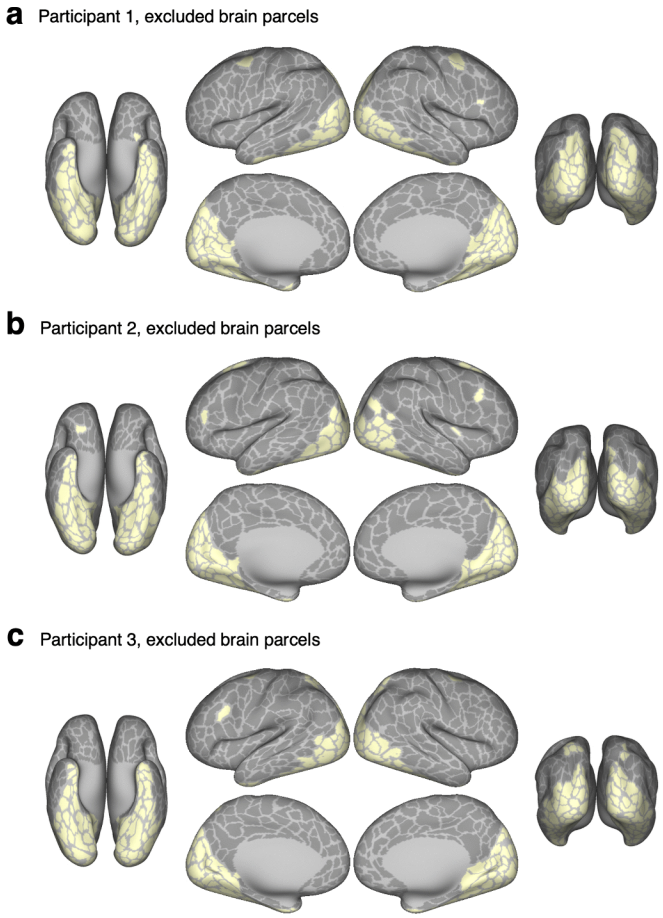
For LASSO, a hyperparameter λ determines the strength of regularization. Within the training data, we used the leave-one-session-out cross-validation to identify the λ that yielded the highest prediction performance at the 1-minute timescale (i.e., nested cross-validation). This procedure separates hyperparameter tuning and model testing into inner and outer loops, providing less biased estimates of prediction performance.



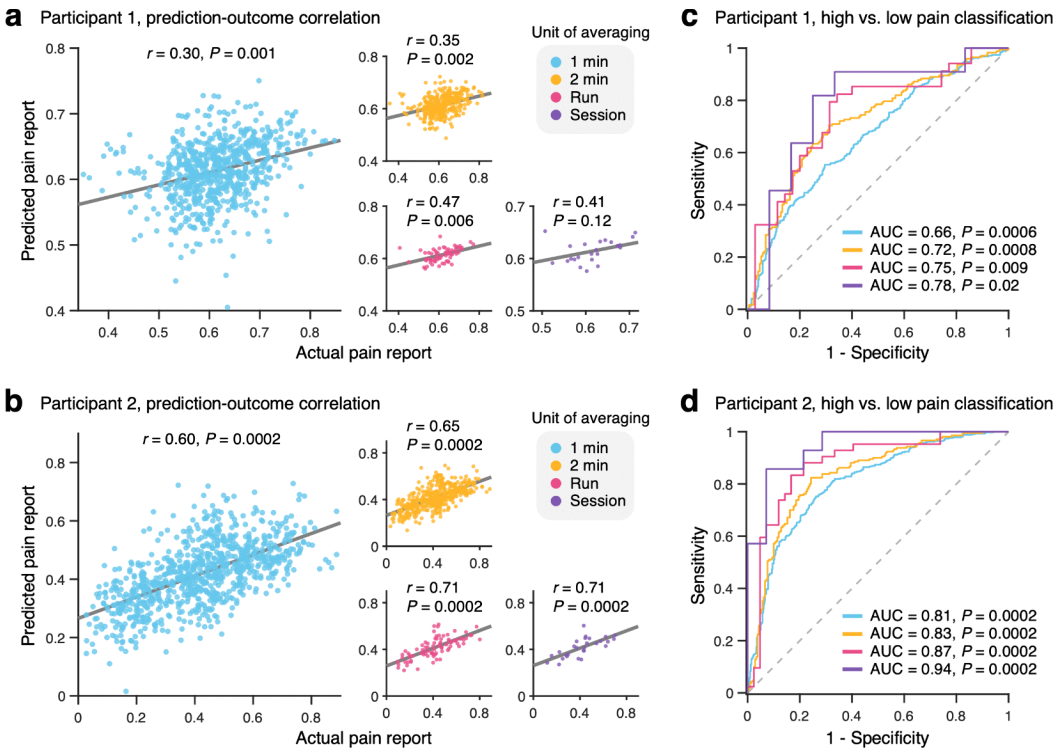
Supplementary Fig. 1. Prediction performance for Participant 3. a, Actual versus predicted pain reports for Participant 3. Colors represent the unit of averaging (i.e., length of time-consecutive bins). Pearson correlations between actual and predicted pain reports are shown in the plots (95% CI: -0.35 to 0.16, -0.44 to 0.26, -0.64 to 0.36, and -0.73 to 0.40 for timescales of 1 min, 2 min, run, and session, respectively). **b,** Receiver operating characteristic (ROC) curves for classifying median-dichotomized high- vs. low-pain states. Colors represent the unit of averaging. Area under the ROC curve (AUC) values are shown in the plots (95% CI: 0.30 to 0.63, 0.26 to 0.67, 0.09 to 0.73, and 0.03 to 0.78 for timescales of 1 min, 2 min, run, and session, respectively).



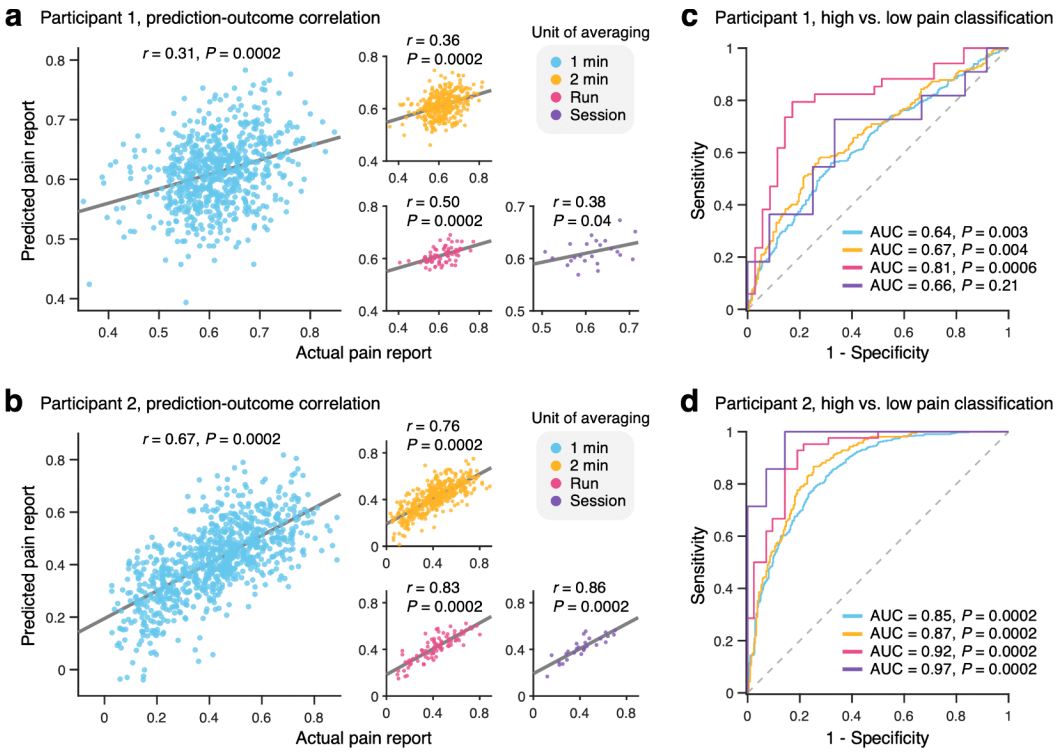
Supplementary Fig. 2. Continuous self-reports of spontaneous pain. Pain intensity ratings over time across all runs and sessions. We excluded the initial 30 seconds of each run, which may be influenced by initial rating-related motion.



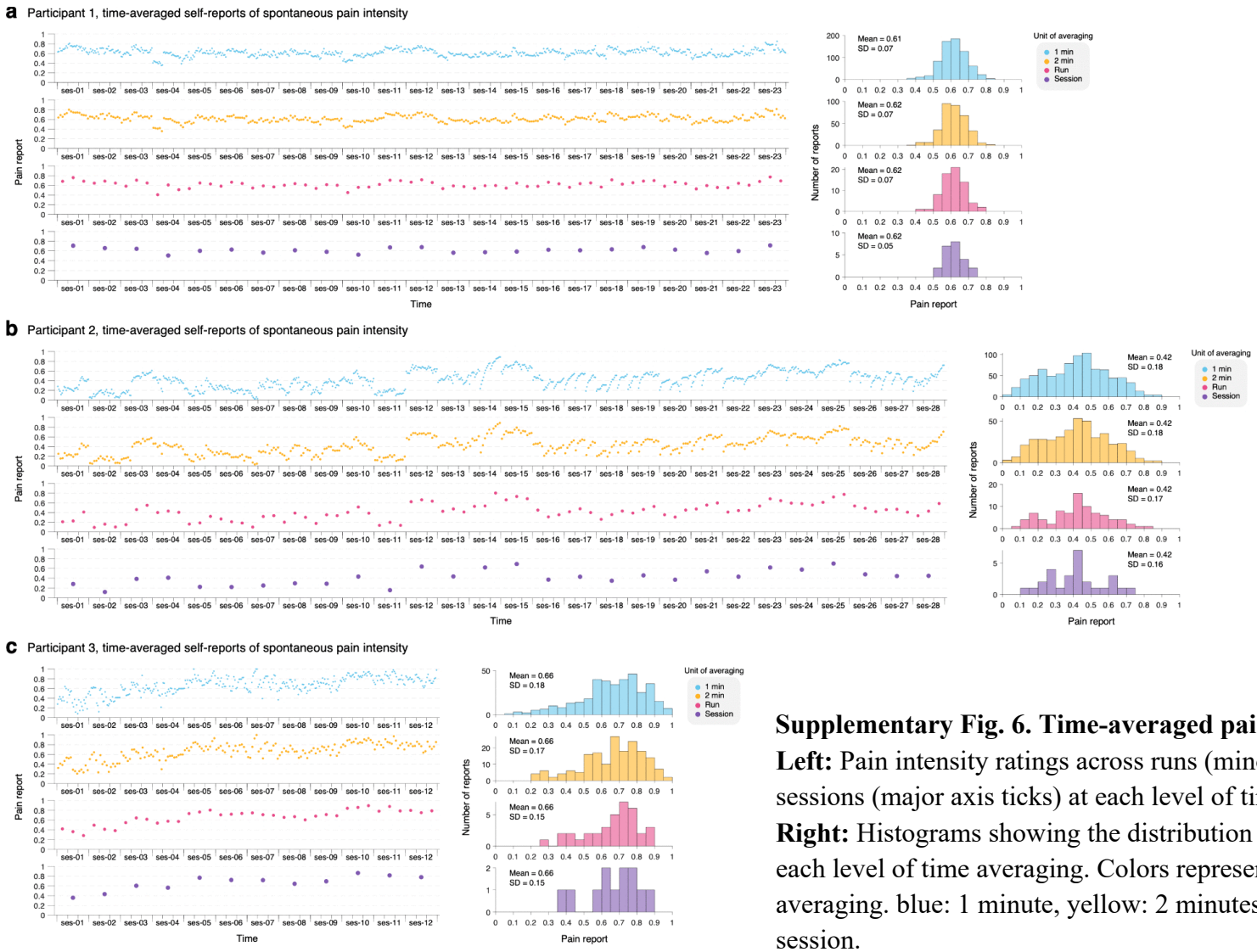
Supplementary Fig. 3. Excluded brain parcels in an individual-specific brain parcellation. Brain parcels likely affected by task-related visual processing were excluded from an individual-specific brain parcellation (highlighted in pale yellow). The number of excluded brain parcels was **(a)** 125 for Participant 1, **(b)** 121 for Participant 2, and **(c)** 120 for Participant 3.



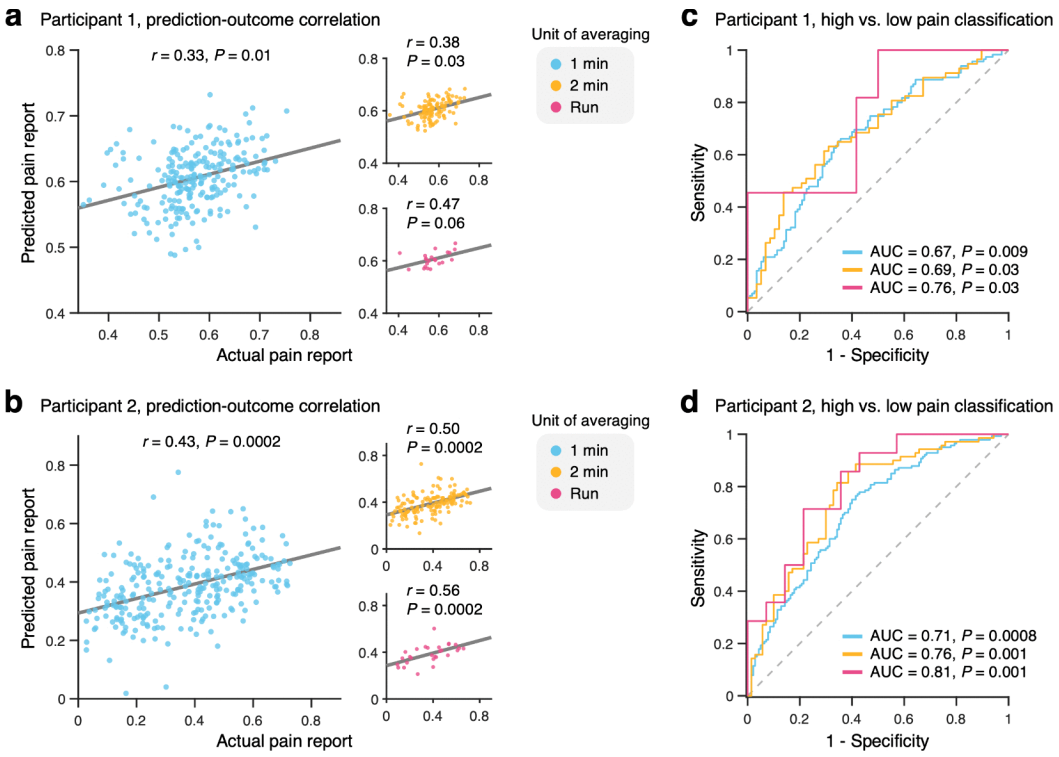
Supplementary Fig. 4. Prediction performance of decoding models including vision-related parcels. We developed personalized decoding models based on the whole brain, including brain parcels related to visual processing (see Supplementary Fig. 3 for details of these regions). **a-b**, Actual versus predicted pain reports for **(a)** Participant 1 and **(b)** Participant 2. Pearson correlations between actual and predicted pain reports are shown in the plots (Participant 1: 95% CI, 0.13 to 0.47, 0.13 to 0.56, 0.12 to 0.78, and -0.10 to 0.83 for timescales of 1 min, 2 min, run, and session, respectively; Participant 2: 95% CI, 0.47 to 0.71, 0.50 to 0.78, 0.54 to 0.85, and 0.52 to 0.86 for the same timescales). **c-d**, ROC curves for classifying median-dichotomized high- vs. low-pain states for **(c)** Participant 1 and **(d)** Participant 2. AUC values are shown in the plots (Participant 1: 95% CI, 0.58 to 0.74, 0.61 to 0.82, 0.57 to 0.91, and 0.55 to 0.97 for timescales of 1 min, 2 min, run, and session, respectively; Participant 2: 95% CI, 0.73 to 0.87, 0.74 to 0.91, 0.76 to 0.96, and 0.85 to 1.00 for the same timescales).



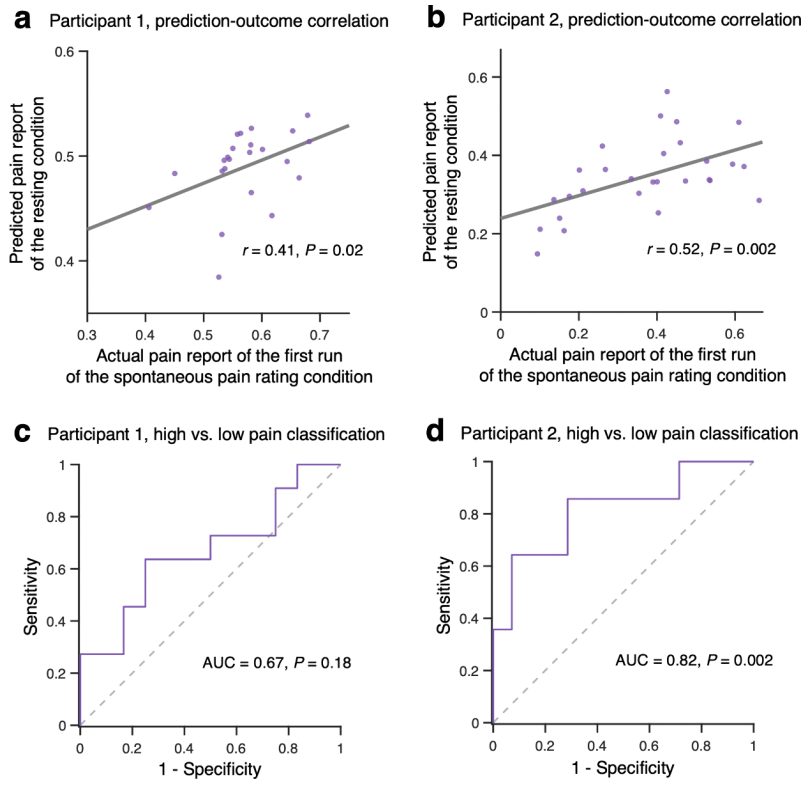
Supplementary Fig. 5. Prediction performance of decoding models using only vision-related parcels. We developed personalized decoding models using only brain parcels related to visual processing (see Supplementary Fig. 3 for details of these regions). **a**-**b**, Actual versus predicted pain reports for **(a)** Participant 1 and **(b)** Participant 2. Pearson correlations between actual and predicted pain reports are shown in the plots (Participant 1: 95% CI, 0.17 to 0.44, 0.19 to 0.51, 0.29 to 0.68, and 0.02 to 0.65 for timescales of 1 min, 2 min, run, and session, respectively; Participant 2: 95% CI, 0.56 to 0.76, 0.65 to 0.84, 0.72 to 0.91, and 0.74 to 0.94 for the same timescales). **c**-**d**, ROC curves for classifying median-dichotomized high- vs. low-pain states for **(c)** Participant 1 and **(d)** Participant 2. AUC values are shown in the plots (Participant 1: 95% CI, 0.55 to 0.72, 0.56 to 0.77, 0.65 to 0.94, and 0.42 to 0.88 for timescales of 1 min, 2 min, run, and session, respectively; Participant 2: 95% CI, 0.77 to 0.90, 0.78 to 0.93, 0.83 to 0.98, and 0.90 to 1.00 for the same timescales).



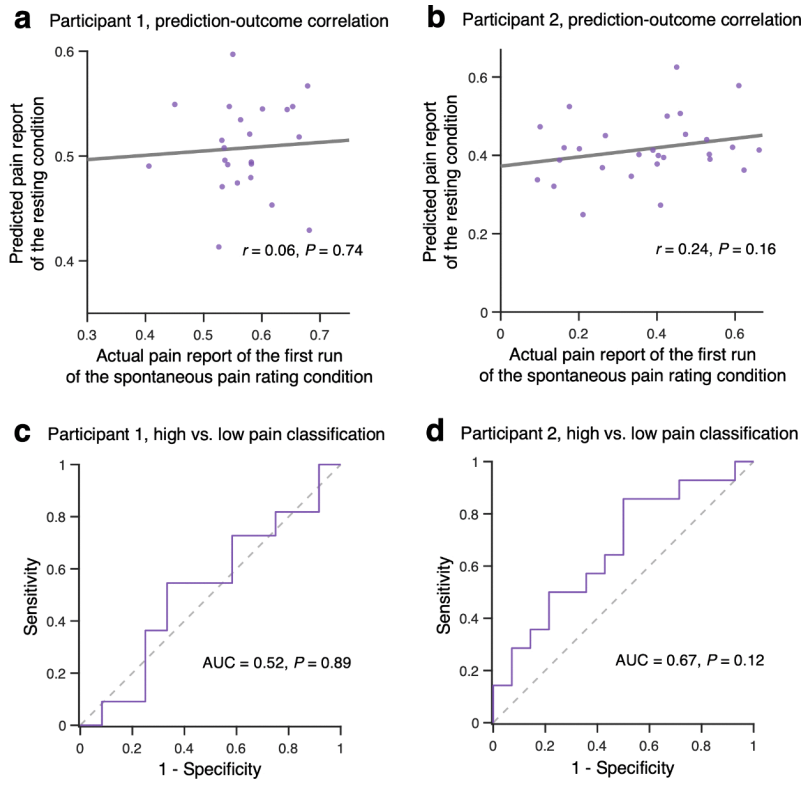
Supplementary Fig. 6. Time-averaged pain intensity ratings. **Left:** Pain intensity ratings across runs (minor axis ticks) and sessions (major axis ticks) at each level of time averaging. **Right:** Histograms showing the distribution of pain ratings at each level of time averaging. Colors represent the unit of averaging. blue: 1 minute, yellow: 2 minutes, red: run, purple: session.



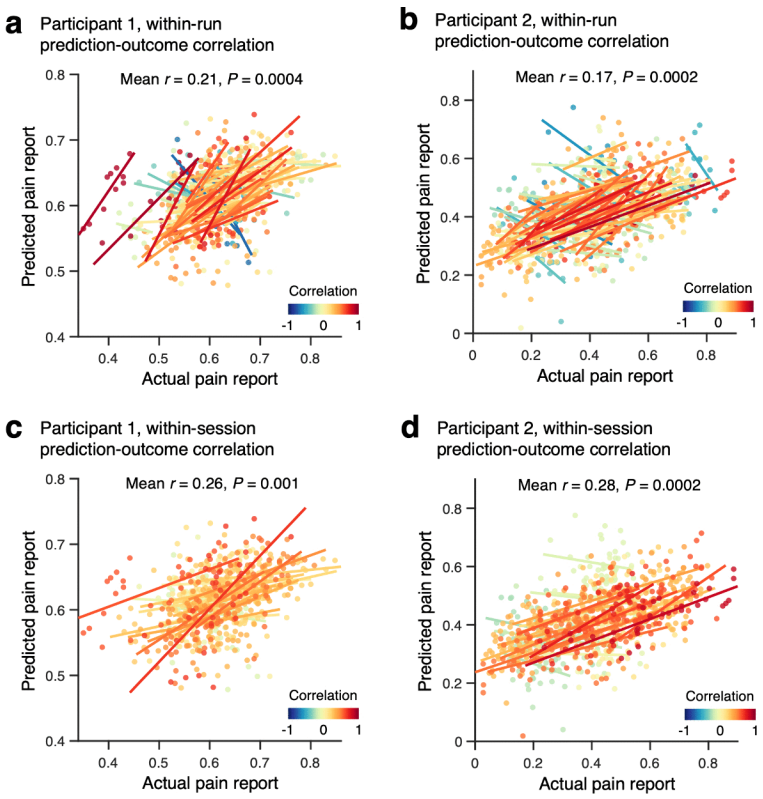
Supplementary Fig. 7. Prediction performance for the first run. We evaluated the personalized decoding models using only the first run data, prior to the pain-exacerbating maneuver. **a**-**b**, Actual versus predicted pain reports for **(a)** Participant 1 and **(b)** Participant 2. Pearson correlations between actual and predicted pain reports are shown in the plots (Participant 1: 95% CI, 0.08 to 0.56, 0.05 to 0.67, and -0.02 to 0.82 for timescales of 1 min, 2 min, and run, respectively; Participant 2: 95% CI, 0.26 to 0.58, 0.28 to 0.68, and 0.36 to 0.77 for the same timescales). **c**-**d**, ROC curves for classifying median-dichotomized high- vs. low-pain states for **(c)** Participant 1 and **(d)** Participant 2. AUC values are shown in the plots (Participant 1: 95% CI, 0.55 to 0.79, 0.52 to 0.84, and 0.52 to 0.93 for timescales of 1 min, 2 min, and run, respectively; Participant 2: 95% CI, 0.59 to 0.82, 0.61 to 0.89, and 0.62 to 0.95 for the same timescales).



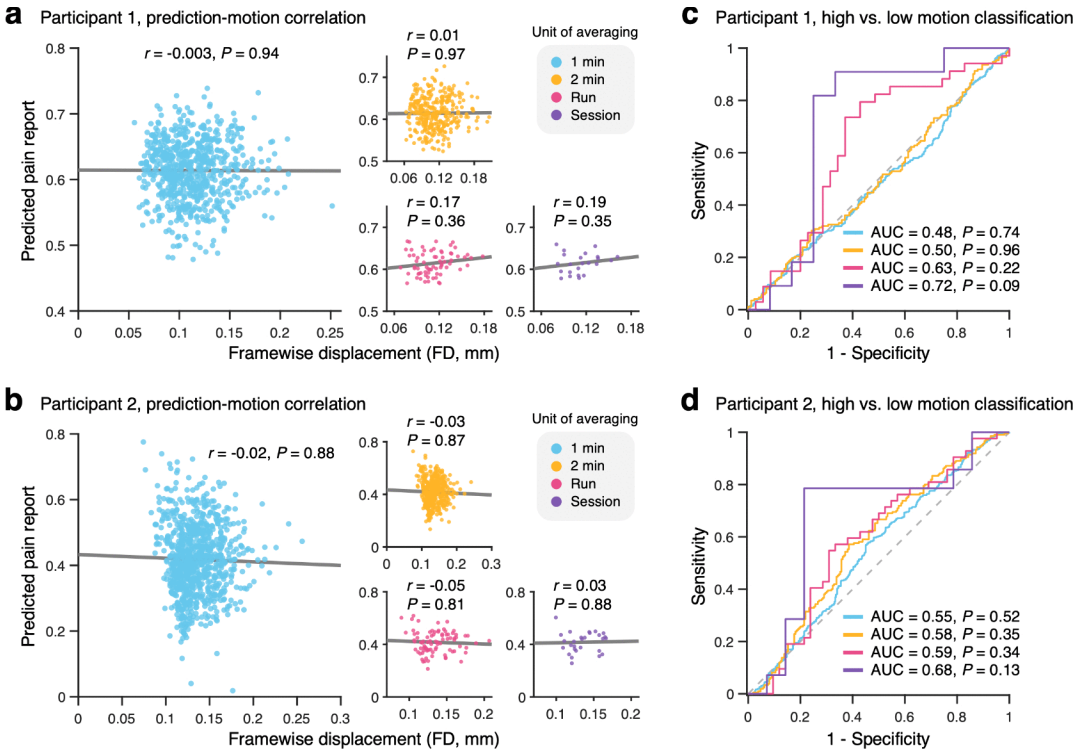
Supplementary Fig. 8. Test results on the resting condition. We tested the personalized decoding models on data from the resting condition to examine whether the models are generalizable to resting-state data. Given that the resting run did not involve any rating procedure and was not used in model training, this test provides a strong test of model generalizability. Model predictions were calculated based on the resting condition and then averaged for each session. These predictions were subsequently compared with the averaged actual pain reports from the first run of the spontaneous pain rating condition, which was closest in time to the resting condition. **a**-**b**, Actual versus predicted pain reports for **(a)** Participant 1 and **(b)** Participant 2. Pearson correlations between actual and predicted pain reports are shown in the plots (Participant 1: 95% CI, 0.10 to 0.66; Participant 2: 95% CI, 0.22 to 0.75). **c**-**d**, ROC curves for classifying median-dichotomized high- vs. low-pain states for **(c)** Participant 1 and **(d)** Participant 2. AUC values are shown in the plots (Participant 1: 95% CI, 0.42 to 0.88; Participant 2: 95% CI, 0.63 to 0.96).



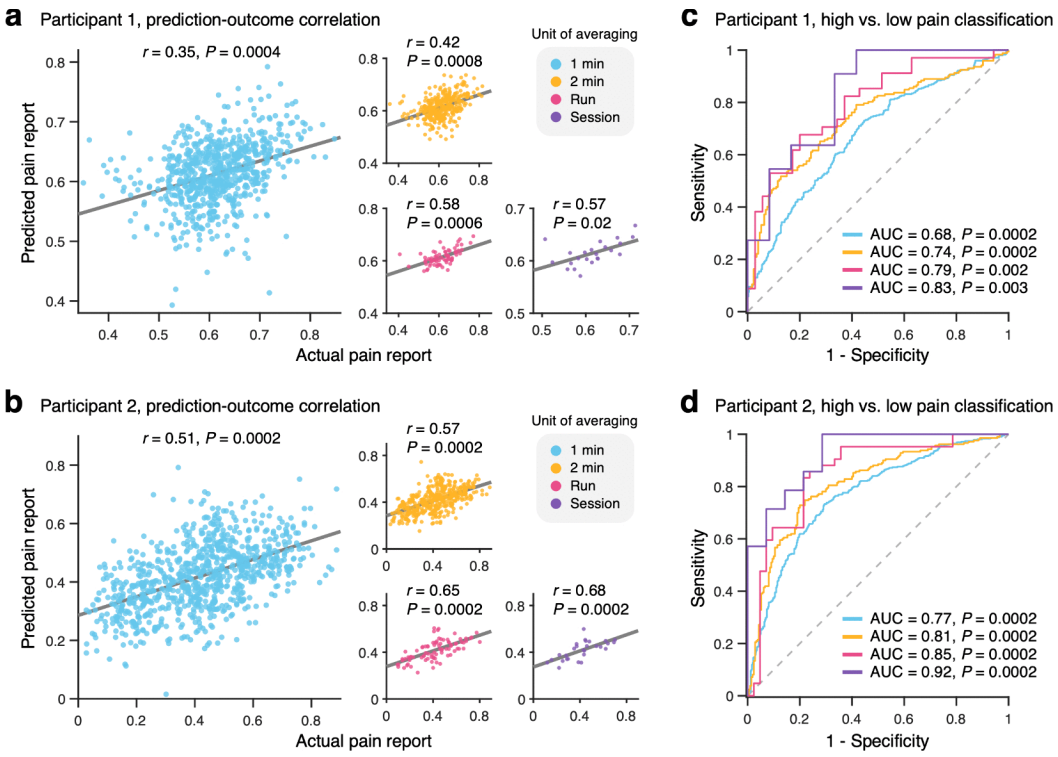
Supplementary Fig. 9. Test results on the resting condition using only vision-related parcels. We tested the personalized decoding models that used only vision-related parcels on data from the resting condition. Model predictions were computed based on the resting-state data and then averaged for each session. These predictions were subsequently compared with the mean actual pain reports from the first run of the spontaneous pain rating condition, which was closest in time to the resting condition. **a**-**b**, Actual versus predicted pain reports for **(a)** Participant 1 and **(b)** Participant 2. Pearson correlations between actual and predicted pain reports are shown in the plots (Participant 1: 95% CI, -0.38 to 0.51; Participant 2: 95% CI, -0.10 to 0.54). **c**-**d**, ROC curves for classifying median-dichotomized high- vs. low-pain states for **(c)** Participant 1 and **(d)** Participant 2. AUC values are shown in the plots (Participant 1: 95% CI, 0.28 to 0.78; Participant 2: 95% CI, 0.45 to 0.86).



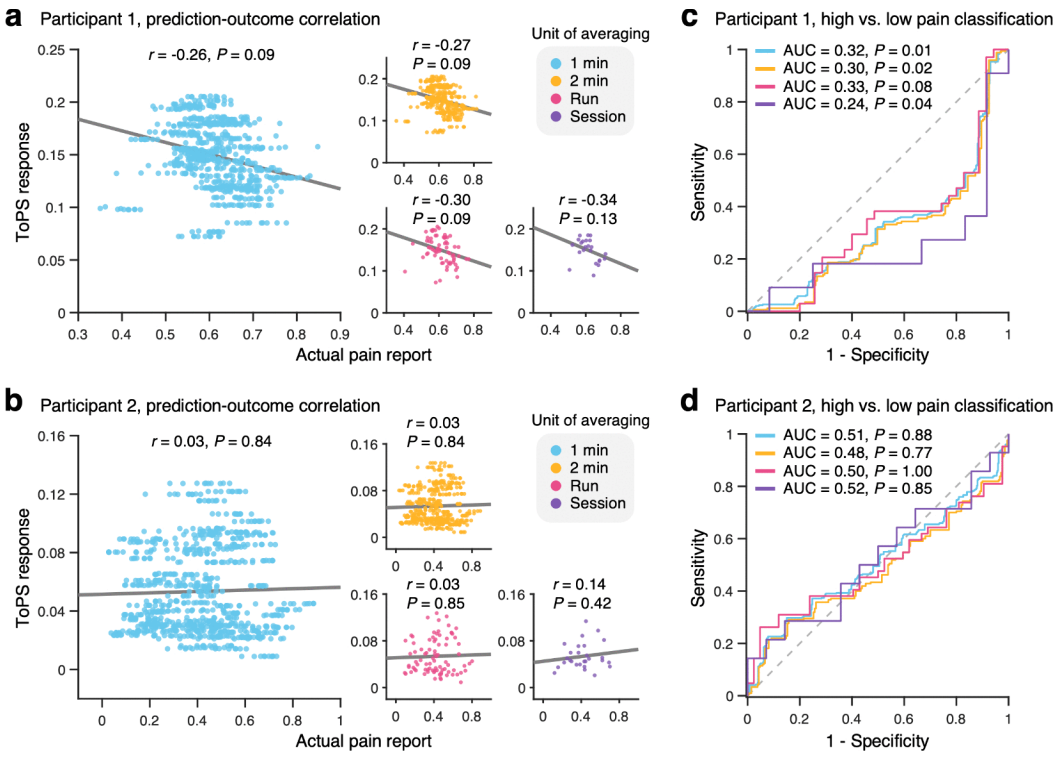
Supplementary Fig. 10. Within-run and within-session prediction performance. We examined whether the personalized decoding models could predict the within-run and within-session variations in spontaneous pain ratings. First, we calculated the Pearson correlation between minute-level (i.e., 10 bins per run) actual and predicted pain reports, either within each run or within each session. Next, we tested whether the distribution of within-run and within-session correlation coefficients was significantly higher than zero using bootstrap tests with 10,000 iterations. **a-b,** Actual versus predicted pain reports for **(a)** Participant 1 and **(b)** Participant 2, with dot and line colors representing within-run correlations. The mean within-run correlations are shown in the plots (Participant 1: 95% CI, 0.10 to 0.32; Participant 2: 95% CI, 0.07 to 0.28). **c-d,** Actual versus predicted pain reports for **(c)** Participant 1 and **(d)** Participant 2, with dot and line colors representing within-session correlations. The mean within-session correlations are shown in the plots (Participant 1: 95% CI, 0.18 to 0.34; Participant 2: 95% CI, 0.17 to 0.38).



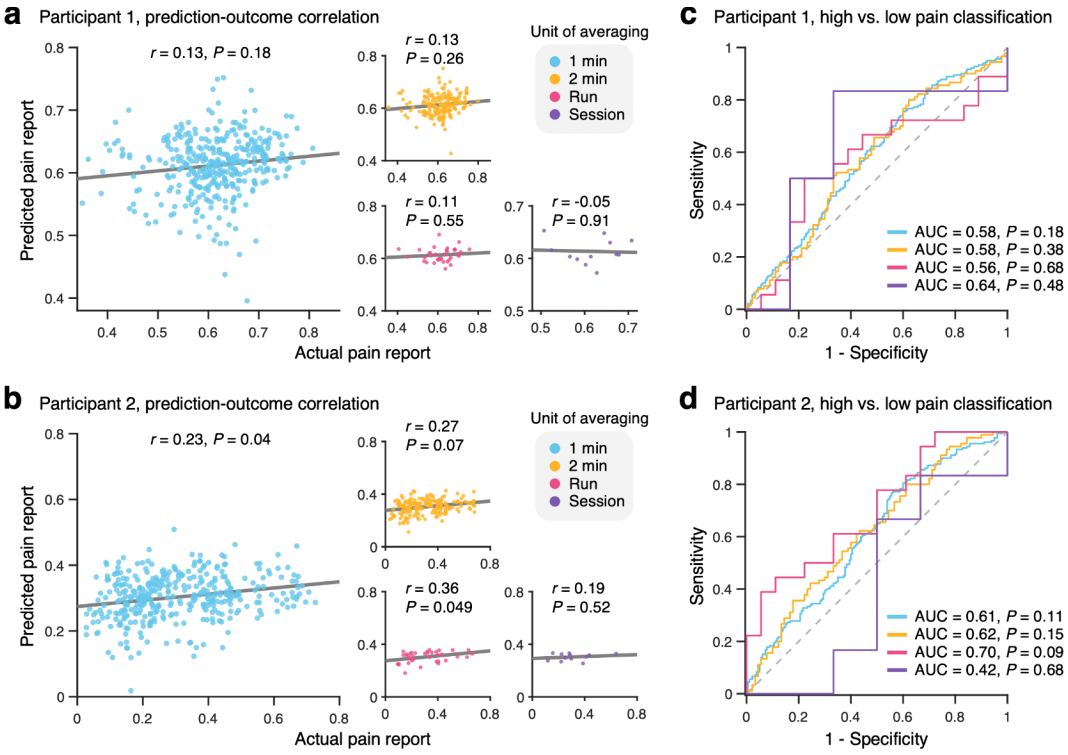
Supplementary Fig. 11. Association between prediction and head motion. We examined whether the predictions from the personalized decoding models were significantly driven by temporal changes in head motion. We used framewise displacement (FD, in mm) as an estimate of head motion. **a-b,** Head motion versus predicted pain reports for **(a)** Participant 1 and **(b)** Participant 2. Colors represent the unit of averaging (i.e., length of time-consecutive bins). Pearson correlations between head motion and predicted pain reports are shown in the plots (Participant 1: 95% CI, -0.19 to 0.17, -0.22 to 0.22, -0.19 to 0.51, and -0.21 to 0.60 for timescales of 1 min, 2 min, run, and session, respectively; Participant 2: 95% CI, -0.27 to 0.23, -0.32 to 0.28, -0.37 to 0.30, and -0.37 to 0.49 for the same timescales). **c-d,** ROC curves for classifying median-dichotomized high- vs. low-head motion states for **(c)** Participant 1 and **(d)** Participant 2. Colors represent the unit of averaging. AUC values are shown in the plots (Participant 1: 95% CI, 0.38 to 0.59, 0.36 to 0.64, 0.42 to 0.84, and 0.47 to 0.94 for timescales of 1 min, 2 min, run, and session, respectively; Participant 2: 95% CI, 0.40 to 0.69, 0.41 to 0.74, 0.40 to 0.78, and 0.45 to 0.89 for the same timescales).



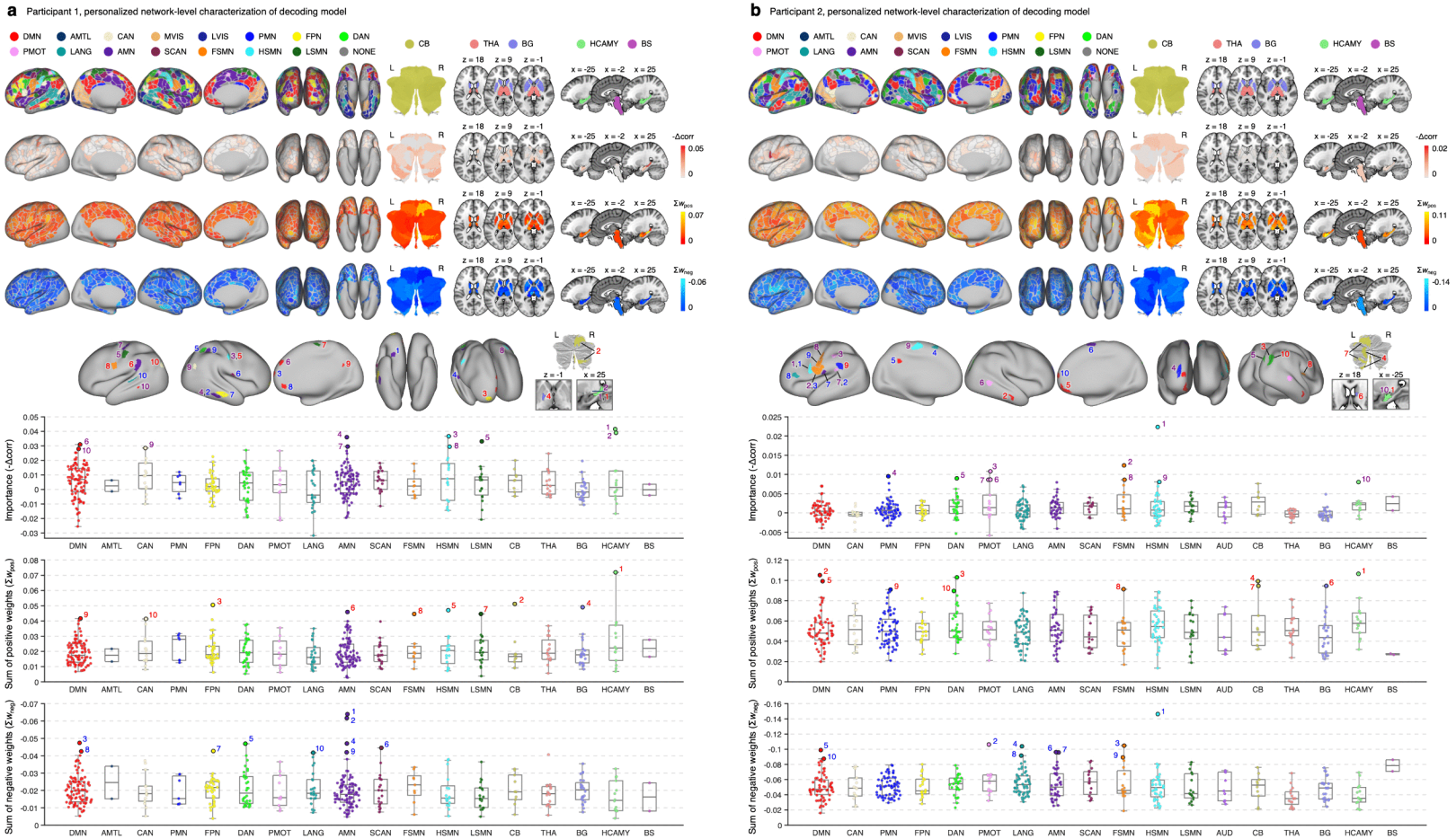
Supplementary Fig. 12. Decoding based on the population-level brain parcellation. We developed personalized decoding models based on a 477-region population-level brain parcellation. This parcellation includes 411 cortical regions from the 500-region Schaefer atlas[16] after removing the visual network, 54 subcortical regions from the Tian atlas[9], 10 cerebellar regions from the multi-domain task battery atlas[10], and the periaqueductal gray and brainstem regions used in previous studies[11,12]. **a-b,** Actual versus predicted pain reports for **(a)** Participant 1 and **(b)** Participant 2. Colors represent the unit of averaging (i.e., length of time-consecutive bins). Pearson's correlations between actual and predicted pain reports are shown in the plots (Participant 1: 95% CI, 0.17 to 0.51, 0.21 to 0.59, 0.27 to 0.81, and 0.10 to 0.88 for timescales of 1 min, 2 min, run, and session, respectively; Participant 2: 95% CI, 0.37 to 0.63, 0.41 to 0.71, 0.48 to 0.81, and 0.50 to 0.84 for the same timescales). **c-d,** ROC curves for classifying median-dichotomized high- vs. low-pain states for **(c)** Participant 1 and **(d)** Participant 2. Colors represent the unit of averaging. AUC values are shown in the plots (Participant 1: 95% CI, 0.59 to 0.76, 0.64 to 0.83, 0.62 to 0.93, and 0.63 to 0.97 for timescales of 1 min, 2 min, run, and session, respectively; Participant 2: 95% CI, 0.68 to 0.84, 0.71 to 0.89, 0.72 to 0.95, and 0.81 to 0.99 for the same timescales).



Supplementary Fig. 13. Testing results of the Tonic Pain Signature (ToPS). We tested the Tonic Pain Signature (ToPS), a previously developed predictive model of sustained pain designed to capture functional connectivity patterns generalizable at the group level. **a**-**b**, Actual versus predicted pain reports for **(a)** Participant 1 and **(b)** Participant 2. Pearson correlations between actual and predicted pain reports are shown in the plots (Participant 1: 95% CI, -0.54 to 0.05, -0.57 to 0.05, -0.61 to 0.05, and -0.75 to 0.10 for timescales of 1 min, 2 min, run, and session, respectively; Participant 2: 95% CI, -0.23 to 0.29, -0.23 to 0.29, -0.25 to 0.31, and -0.21 to 0.47 for the same timescales). **c**-**d**, ROC curves for classifying median-dichotomized high- vs. low-pain states for **(c)** Participant 1 and **(d)** Participant 2. AUC values are shown in the plots (Participant 1: 95% CI, 0.20 to 0.46, 0.16 to 0.46, 0.16 to 0.52, and 0.05 to 0.48 for timescales of 1 min, 2 min, run, and session, respectively; Participant 2: 95% CI, 0.37 to 0.65, 0.33 to 0.63, 0.33 to 0.67, and 0.29 to 0.74 for the same timescales).

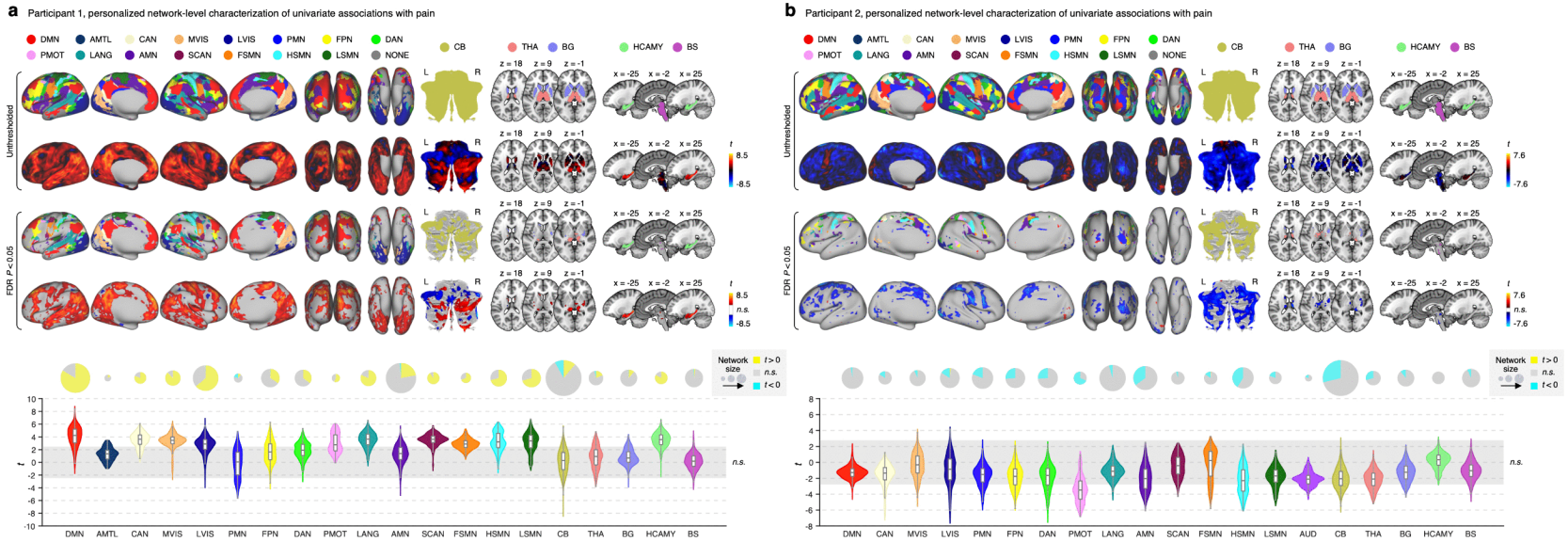


Supplementary Fig. 14. Decoding only with the first 12 sessions. We developed personalized decoding models using the first 12 sessions to match the training data size of Participant 3. **a-b**, Actual versus predicted pain reports for **(a)** Participant 1 and **(b)** Participant 2. Pearson correlations between actual and predicted pain reports are shown in the plots (Participant 1: 95% CI, -0.06 to 0.36, -0.11 to 0.43, -0.27 to 0.60, and -0.59 to 0.63 for timescales of 1 min, 2 min, run, and session, respectively; Participant 2: 95% CI, 0.01 to 0.41, -0.02 to 0.50, 0.003 to 0.64, and -0.51 to 0.69 for the same timescales). **c-d**, ROC curves for classifying median-dichotomized high- vs. low-pain states for **(c)** Participant 1 and **(d)** Participant 2. AUC values are shown in the plots (Participant 1: 95% CI, 0.46 to 0.70, 0.41 to 0.74, 0.27 to 0.84, and 0.26 to 1.00 for timescales of 1 min, 2 min, run, and session, respectively; Participant 2: 95% CI, 0.48 to 0.74, 0.45 to 0.79, 0.46 to 0.91, and 0.06 to 0.80 for the same timescales).



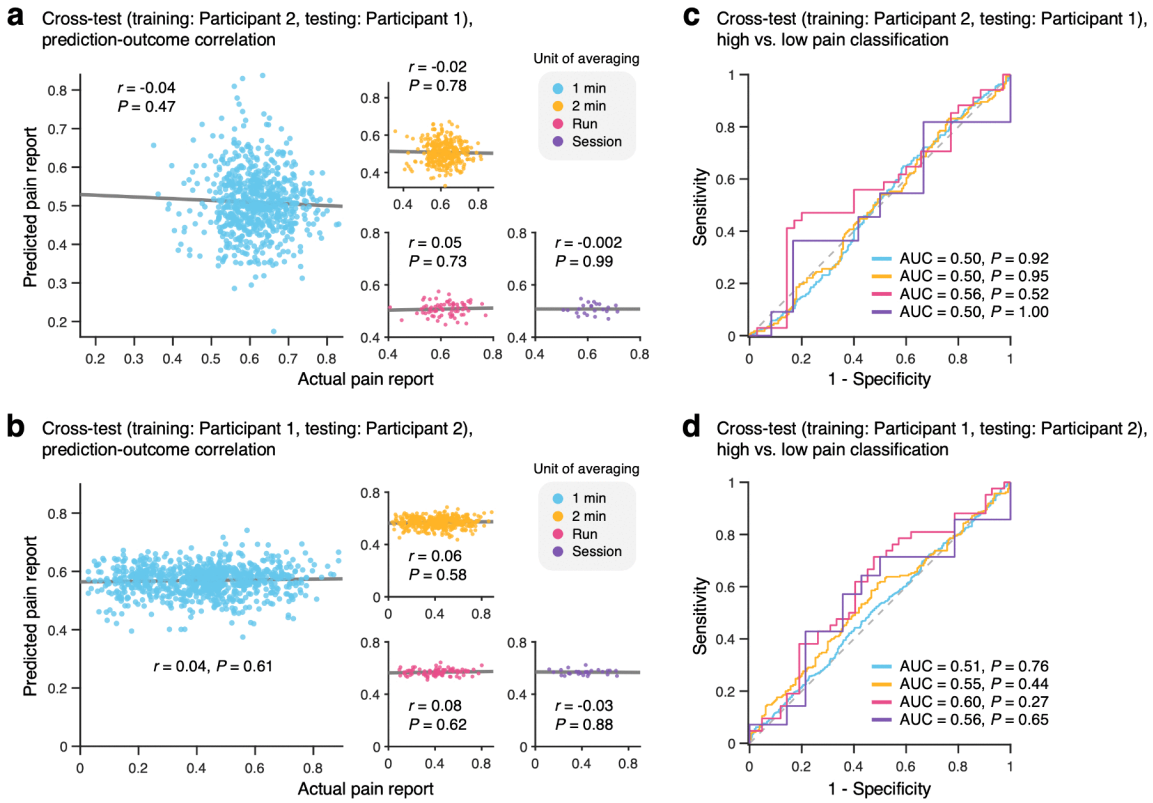
Supplementary Fig. 15. Personalized network-level characterization of the decoding models of spontaneous pain. In the top panel, the first row represents parcel-wise personalized network assignments (see Supplementary Methods), and the second, third, and fourth rows display permutation feature importance (-Δcorr), sum of positive model weights (Σw_{pos}), and sum of negative model weights (Σw_{neg}), respectively. In the bottom panel, box plots depict the distribution of importance (top), sum of positive weights (middle), and sum of negative weights (bottom) for each network. Each box spans the first to the third quartile, with the horizontal line inside the box representing the median. Whiskers extend to the smallest and largest values within 1.5 times the interquartile range

from the lower and upper quartiles. Dots represent individual parcels. The top 10 parcels ranked by importance (purple), sum of positive weights (red), and sum of negative weights (blue) are annotated with numbered labels in the box plots, and their corresponding brain locations are shown in the brain. Abbreviations: DMN, default mode network; AMTL, anterior medial temporal network; CAN, contextual association network; MVIS, medial visual network; LVIS, lateral visual network; PMN, parietal memory network; FPN, frontoparietal network; DAN, dorsal attention network; PMOT, premotor network; LANG, language network; AMN, action mode network; SCAN, somato-cognitive action network; FSMN, face somatomotor network; HSMN, hand somatomotor network; LSMN, leg somatomotor network; AUD, auditory network; CB, cerebellum; THA, thalamus; BG, basal ganglia; HCAMY, hippocampus/amygdala; BS, brainstem.



Supplementary Fig. 16. Personalized network-level characterization of univariate associations with spontaneous pain ratings. We examined univariate vertex- and voxel-wise associations with pain using a general linear model (GLM) analysis for **(A)** Participant 1 and **(B)** Participant 2. For each session, we regressed fMRI activity of every vertex or voxel (Y) on pain ratings convolved with the canonical hemodynamic response function (X). Using the beta maps from the first-level GLM analysis, we performed one-sample *t*-tests, treating session as a random effect. In the top panel, the first row represents vertex- and voxel-wise personalized network assignments (see Supplementary Methods), and the second row represents *t*-values. The third and fourth rows represent personalized network assignments and *t*-values for vertices and voxels that survived false discovery rate (FDR)-corrected thresholding at $P < 0.05$. In the bottom panel, violin plots represent the distribution of *t*-values for each network, with gray shading indicating the range of non-significant (*n.s.*) values. Pie charts represent the proportions of positive (yellow), non-significant (gray), and negative (cyan) vertices and voxels within each network, with circle sizes indicating the number of vertices or voxels in the network. Abbreviations: DMN, default mode network; AMTL, anterior medial temporal network; CAN, contextual association network; MVIS, medial visual network; LVIS, lateral visual network; PMN, parietal memory network; FPN, frontoparietal network; DAN, dorsal attention network; PMOT, premotor network; LANG, language network; AMN, action mode network; SCAN, somato-

cognitive action network; FSMN, face somatomotor network; HSMN, hand somatomotor network; LSMN, leg somatomotor network; AUD, auditory network; CB, cerebellum; THA, thalamus; BG, basal ganglia; HCAMY, hippocampus/amygdala; BS, brainstem.



Supplementary Fig. 17. Cross-testing results based on the population-level brain parcellation. Cross-testing of the personalized decoding models based on a 477-region population-level brain parcellation. **a-b,** Actual versus predicted pain reports for **(a)** Participant 1 and **(b)** Participant 2. Colors represent the unit of averaging. Pearson correlations between actual and predicted pain reports are shown in the plots (Participant 1: 95% CI, -0.16 to 0.07, -0.19 to 0.14, -0.24 to 0.35, and -0.34 to 0.36 for timescales of 1 min, 2 min, run, and session, respectively; Participant 2: 95% CI, -0.12 to 0.21, -0.15 to 0.27, -0.24 to 0.41, and -0.39 to 0.37 for the same timescales). **c-d,** ROC curves for classifying median-dichotomized high- vs. low-pain states for **(c)** Participant 1 and **(d)** Participant 2. Colors represent the unit of averaging. AUC values are shown in the plots (Participant 1: 95% CI, 0.42 to 0.57, 0.40 to 0.59, 0.37 to 0.75, and 0.25 to 0.75 for timescales of 1 min, 2 min, run, and session, respectively; Participant 2: 95% CI, 0.42 to 0.60, 0.43 to 0.66, 0.42 to 0.77, and 0.33 to 0.79 for the same timescales).

Supplementary Table 1. Prediction performance

Participant	Unit of averaging	*N*	Prediction of spontaneous pain report				Classification between high vs. low pain	
			Correlation		R^2	MAE	AUC (95% CI)	*P*
			r (95% CI)	*P*				
Participant 1	1 min	690	0.40 (0.23-0.55)	< 0.001	0.11	0.05	0.71 (0.62-0.79)	< 0.001
	2 min	345	0.47 (0.27-0.65)	< 0.001	0.22	0.05	0.76 (0.65-0.86)	< 0.001
	Run	69	0.59 (0.29-0.82)	< 0.001	0.32	0.04	0.81 (0.63-0.94)	< 0.001
	Session	23	0.61 (0.17-0.89)	0.01	0.35	0.03	0.87 (0.69-1.00)	0.002
Participant 2	1 min	840	0.51 (0.37-0.63)	< 0.001	0.25	0.13	0.76 (0.68-0.84)	< 0.001
	2 min	420	0.56 (0.40-0.70)	< 0.001	0.32	0.12	0.80 (0.71-0.88)	< 0.001
	Run	84	0.63 (0.45-0.79)	< 0.001	0.38	0.11	0.84 (0.70-0.94)	< 0.001
	Session	28	0.65 (0.46-0.82)	< 0.001	0.39	0.10	0.93 (0.82-1.00)	< 0.001

Note. CI, confidence interval; MAE, mean absolute error; AUC, area under the curve.

SUPPLEMENTARY REFERENCES

1. Dickie, E.W., *et al.* Ciftify: A framework for surface-based analysis of legacy MR acquisitions. *Neuroimage* **197**, 818-826 (2019).
2. Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L. & Petersen, S.E. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* **59**, 2142-2154 (2012).
3. Fair, D.A., *et al.* Correction of respiratory artifacts in MRI head motion estimates. *Neuroimage* **208**, 116400 (2020).
4. Newbold, D.J., *et al.* Plasticity and Spontaneous Activity Pulses in Disused Human Brain Circuits. *Neuron* **107**, 580-589 e586 (2020).
5. Behzadi, Y., Restom, K., Liau, J. & Liu, T.T. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage* **37**, 90-101 (2007).
6. Kim, J., *et al.* Somatotopically specific primary somatosensory connectivity to salience and default mode networks encodes clinical pain. *Pain* **160**, 1594-1605 (2019).
7. Gordon, E.M., *et al.* Precision Functional Mapping of Individual Human Brains. *Neuron* **95**, 791-807 e797 (2017).
8. Desikan, R.S., *et al.* An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**, 968-980 (2006).
9. Tian, Y., Margulies, D.S., Breakspear, M. & Zalesky, A. Topographic organization of the human subcortex unveiled with functional connectivity gradients. *Nat Neurosci* **23**, 1421-1432 (2020).
10. King, M., Hernandez-Castillo, C.R., Poldrack, R.A., Ivry, R.B. & Diedrichsen, J. Functional boundaries in the human cerebellum revealed by a multi-domain task battery. *Nat Neurosci* **22**, 1371-1378 (2019).
11. Roy, M., *et al.* Representation of aversive prediction errors in the human periaqueductal gray. *Nat Neurosci* **17**, 1607-1612 (2014).
12. Beissner, F., Schumann, A., Brunn, F., Eisentrager, D. & Bar, K.J. Advances in functional magnetic resonance imaging of the human brainstem. *Neuroimage* **86**, 91-98 (2014).
13. Rosvall, M. & Bergstrom, C.T. Maps of random walks on complex networks reveal community structure. *Proc Natl Acad Sci U S A* **105**, 1118-1123 (2008).
14. Gordon, E.M., *et al.* A somato-cognitive action network alternates with effector regions in motor cortex. *Nature* **617**, 351-359 (2023).
15. Gordon, E.M., *et al.* Default-mode network streams for coupling to language and control systems. *Proc Natl Acad Sci U S A* **117**, 17308-17319 (2020).
16. Schaefer, A., *et al.* Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb Cortex* **28**, 3095-3114 (2018).