nature neuroscience

Article

https://doi.org/10.1038/s41593-023-01338-z

First-in-human prediction of chronic pain state using intracranial neural biomarkers

In the format provided by the authors and unedited

nature neuroscience

Article

https://doi.org/10.1038/s41593-023-01338-z

First-in-human prediction of chronic pain state using intracranial neural biomarkers

In the format provided by the authors and unedited

Supplementary Material

Table of Contents

Supplementary Methods

Justification for Dichotomizing Pain Scores

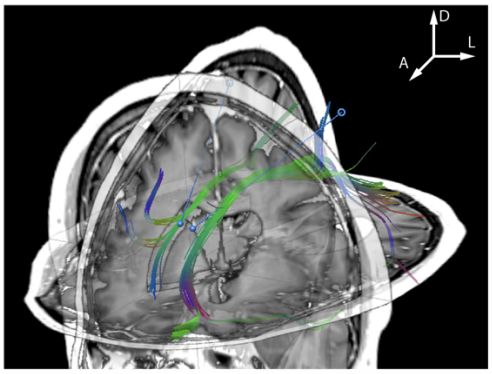
Pain score outcome data were dichotomized because one main research goal was to assess the performance of such dichotomized data in the real-world, when implemented in a closed-loop brain stimulation device that requires defining binary symptom states. The device used for recording neural activity in this study, Medtronic Activa PC+S, is also capable of adaptively adjusting brain stimulation when the output of a linear discriminant crosses a specified threshold from one binary state (e.g. low pain) to another (e.g. high pain). A newly released commercial DBS system, Medtronic Percept, follows this same framework. The linear discriminant control signal is computed as a function of selected power band features; therefore, one main goal of the study was to determine which power band features would be most predictive of dichotomized pain state to inform a future closed-loop control signal in each subject. Therefore, when designing the study, we sought to undertake analysis that closely matched the demands of available clinical technology for ease of interpretation and feasibility of future closed-loop stimulation control.

DeCoster et al.,[1,2] have provided a useful framework to guide the artificial dichotomization of data. They argue that dichotomization is specifically acceptable when three criteria are met 1) when the underlying variable is naturally categorical (such as NRS), 2) the observed measure has high reliability and 3) the relative group sizes of the dichotomized indicator match those of the underlying variable. The dichotomized variable (pain NRS) in the present study meets all three criteria. Specifically, the underlying variable (NRS) is itself naturally categorical and has high reliability (as evidenced high correlation with the separately reported variable of pain VAS in **Figure 1D**). Further, the group sizes of the dichotomized indicator are roughly equal in most cases (based on median split), and visually match the relative distribution of low vs high pain scores. Most importantly, the pain score dichotomization matches the intuitive grouping of the reported NRS scores as ‘high’ or ‘low’ by each subject, when they were queried about the median split used in our analyses.

Farrington and Loeber[3] further highlight that dichotomization is more appropriate when distribution are very skewed, as is the case in all subjects in this manuscript that exhibited much more frequent high pain scores than medium or low scores both in the acute and chronic settings. Further, dichotomization may be beneficial when a variable is not linearly related to the measured outcome; in our case only a handful of band power variables exhibited a linear relationship with pain score, with high pain scores being associated with both high and lower power in most bands.

Supplementary Figures and Results

A.



B.

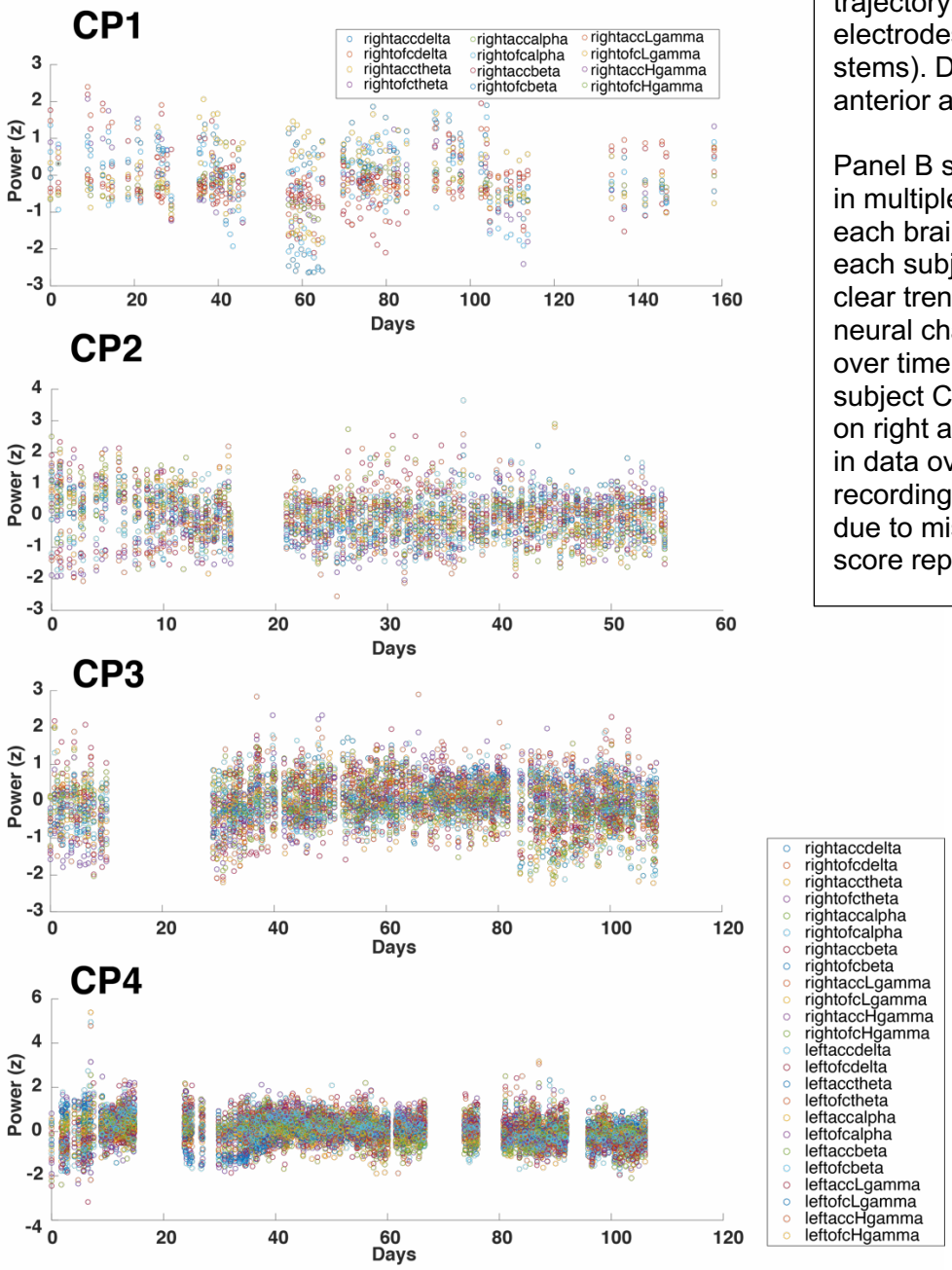


Figure S1. DTI based electrode targeting and signal stability

Panel A shows a Diffusion Tensor Image tracing of the cingulum bundle bilaterally in one subject (green fibers represent those running in anterior-posterior direction) and example trajectory of bilateral depth electrodes targeting ACC (blue stems). D,A,L represent dorsal, anterior and lateral, respectively.

Panel B shows normalized power in multiple frequency bands from each brain region over time for each subject. There were no clear trends that indicated non-neural changes in recordings over time. Top legend applies to subject CP1 while legend in box on right applies to CP2-4. Gaps in data over time indicate missing recordings or those eliminated due to misalignment with pain score reporting.

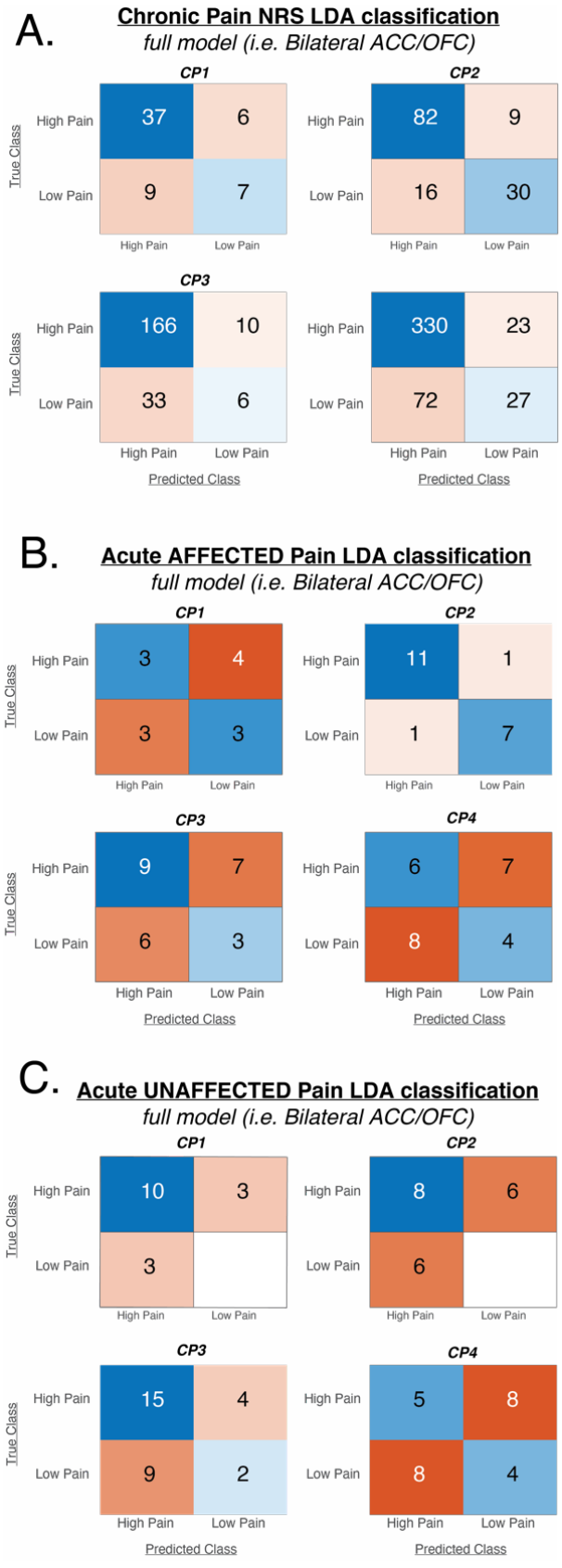


Figure S2. Confusion Matrices of LDA Full Models.

Panels A-C show LDA classification tables with raw counts of all true versus predicted samples for all full models of spontaneous, chronic pain (A), and acute pain on the affected side (B) and unaffected side (C). Box color intensity represents proportion of total samples in blue for accurate classifications and red for false positives or negatives.

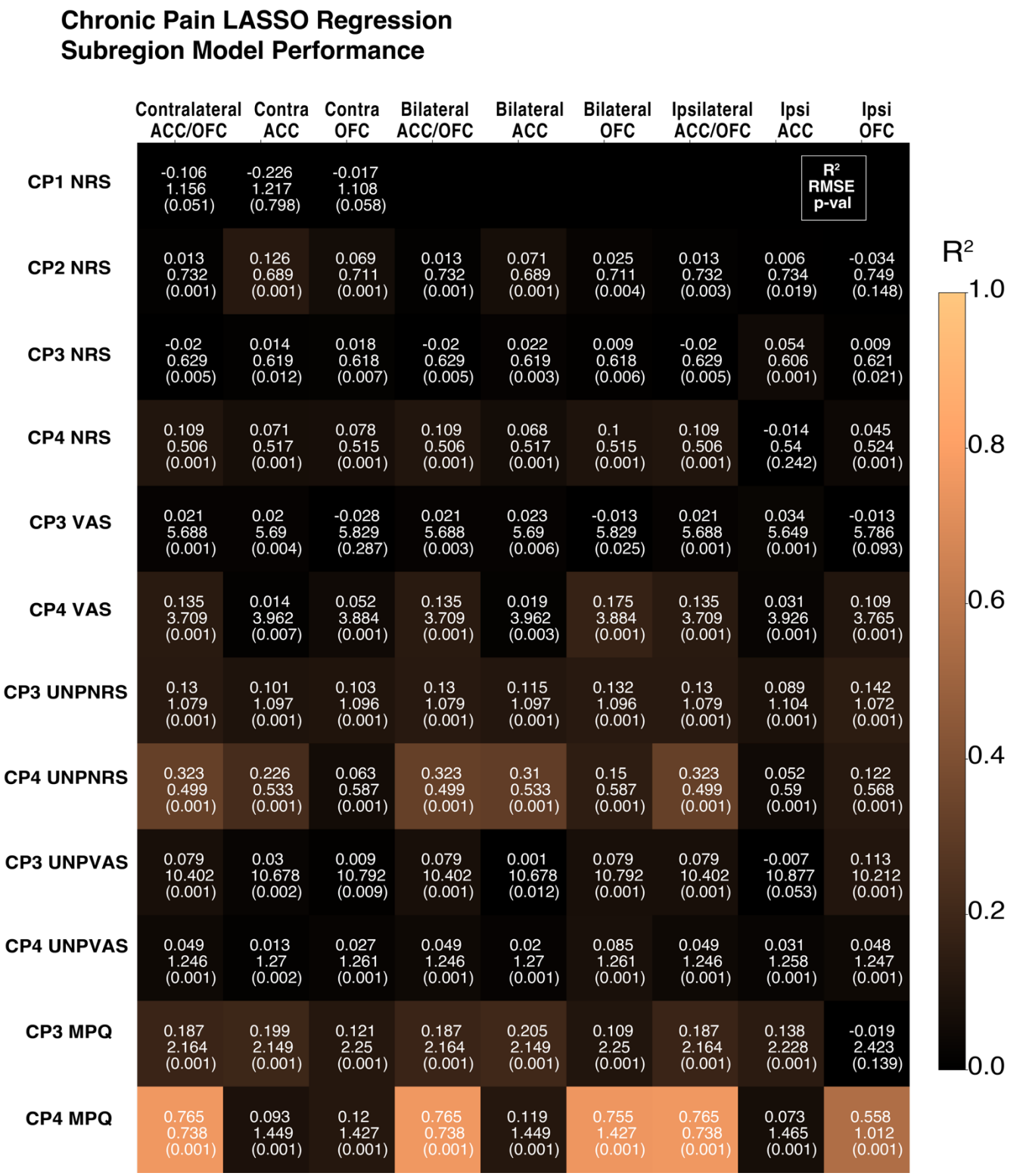


Figure S3. Chronic Pain Subregion Regression Model Performance (LASSO regression)

Heatmap table shows coefficient of determinations (R^2), root mean squared error (RMSE), and one-sided p-values from permutation tests (n=1000, without multiple comparisons adjustment) in parenthesis for all calculated chronic pain models (across columns) and pain score report types (along rows) for regression. R^2 also represented in color based on color bar scale on right. The only R^2 values > 0.3 occurred for CP4 Unpleasantness NRS for models including OFC+ACC and for SF-MPQ, when OFC was included in models.

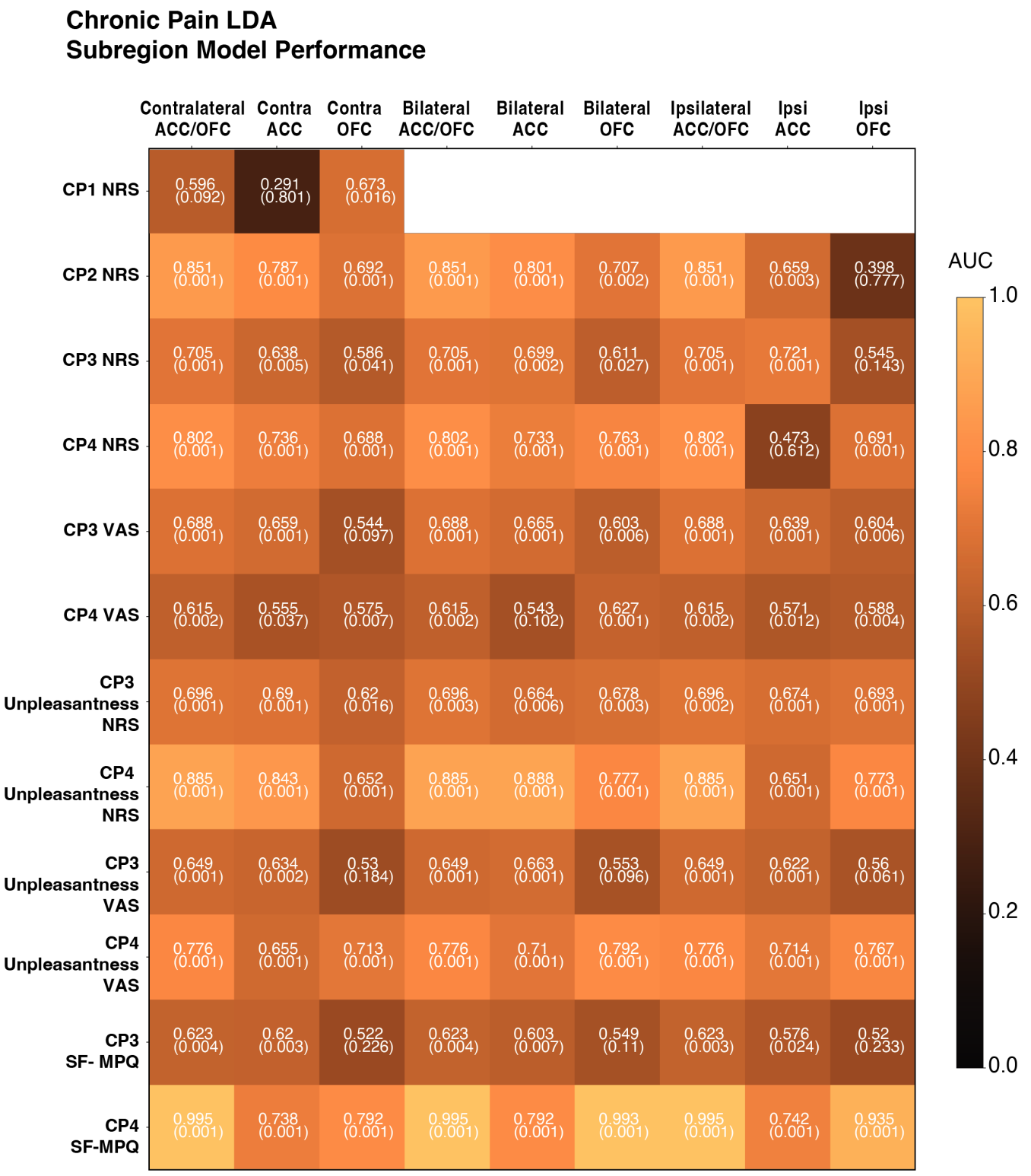


Figure S4. Chronic Pain Subregion Classification Model Performance (LDA)

Heatmap table shows AUC and one-sided empirical p-values based on permutation tests (n=1000, without multiple comparisons adjustment) below in parenthesis for all cross-validated chronic pain models (across columns) and pain score report types (across rows) underlying Figure 2. AUC also represented in color based on color bar scale on right.

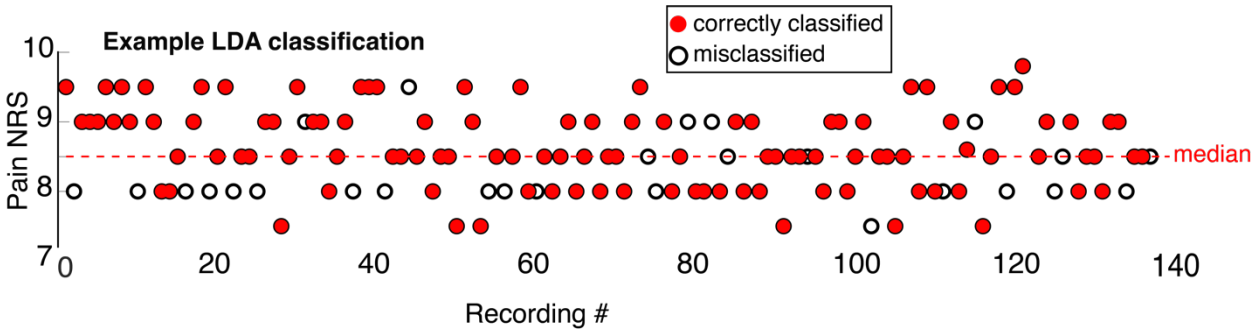


Figure S5. Example LDA classification performance at individual recording session level in one subject.

Example of classifier performance over recordings using LDA with leave one out cross validation for CP2 pain NRS full model. Filled red circles indicate successful classification, while empty circles indicate misclassified pain reports. Dashed horizontal line shows median NRS value for this example.

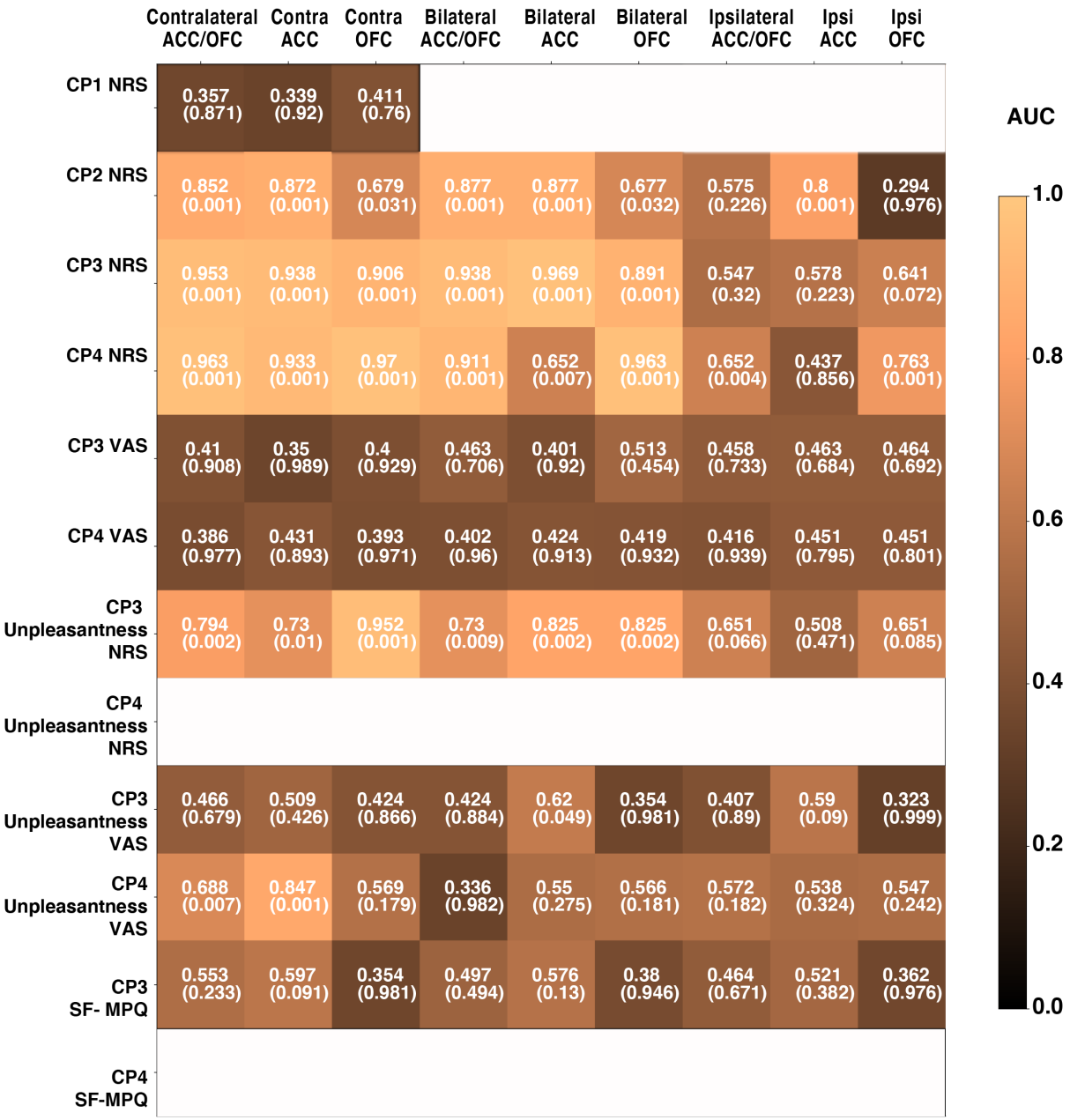


Figure S6. Chronic pain subregion Model Stability over time (LDA) using 70/30 split training/ testing datasets.

Heatmap table shows AUC and one-sided empirical p-values based on permutation tests (n=1000, without multiple comparisons adjustment) below in parenthesis for all calculated chronic pain models using the first 70% of data for training and the last 30% for testing (across columns) and pain score report types (across rows). For pain intensity NRS, all models were similarly significant to Figure S4 except for CP1 contraOFC and CP2 IpsiACC/OFC. Note that pain intensity VAS for CP3 and CP4 were not significant (when using 70/30 split method) due to increased variance in the latter half of the data reflecting decreased consistency in VAS reporting during this period. Missing data for CP4 indicates occasions where the last 30% of data contained no variation in pain metric values so models could not be computed.

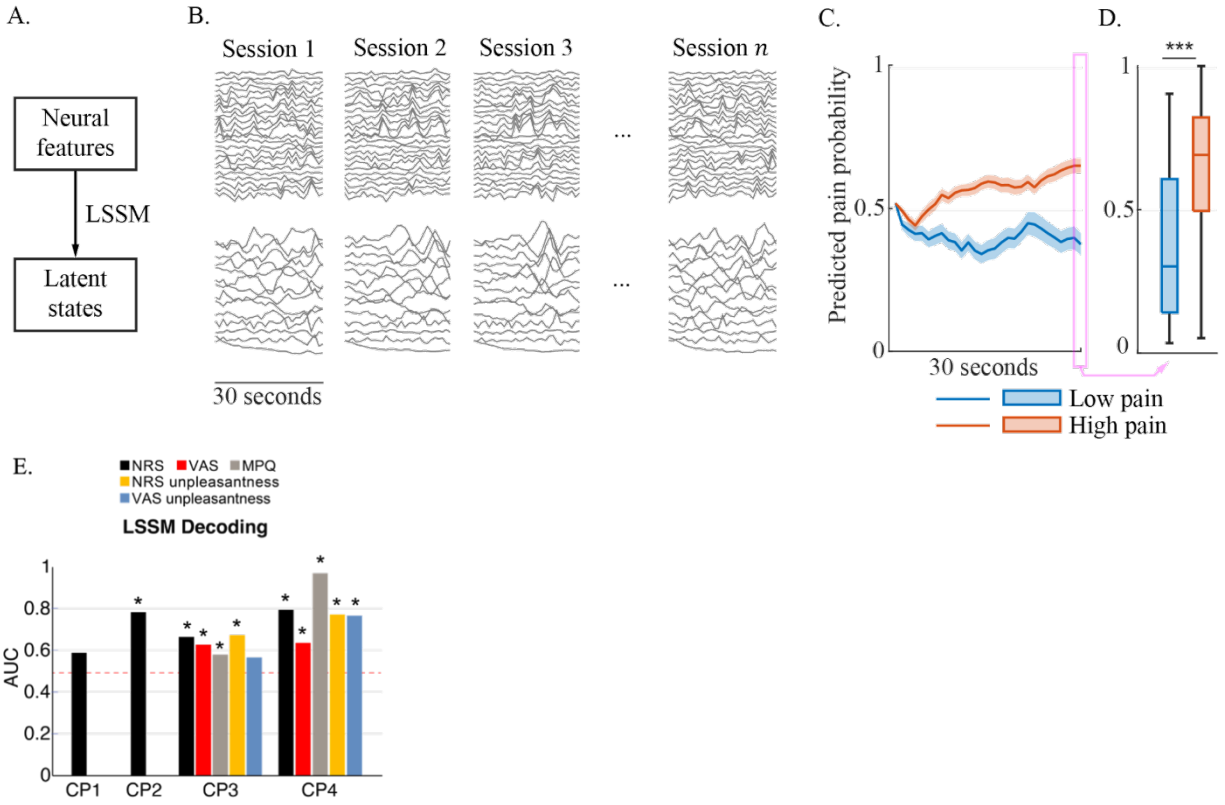


Figure S7: LSSM Schematic: Dynamic models can decode chronic pain state from neural activity.

A) The decoder consists of a dynamic linear state space model (LSSM) that summarizes the information in neural features into a latent state, which is then provided to a classifier to estimate the probability of high chronic pain. Neural features consisted of log-powers computed in canonical frequency bands in non-overlapping 1-second time windows (Methods). B) Example time series for the neural features (first row) and the extracted latent states (second row) are shown for several recording sessions in subject CP2. C) Predicted probability of the high pain class through the length of a recording session, averaged over sessions with high pain (red, $\mathrm{N} = 91$), and over session with low pain (blue, $\mathrm{N} = 46$). Shaded areas show the S.E.M. D) Extract of data shown in C, showing only the last time step of the recording session, which is used for final decoding. Predicted pain probability was significantly higher for the $\mathrm{N} = 91$ sessions that had high pain compared to the $\mathrm{N} = 46$ sessions that had low pain (***$\mathrm{P} = 3.7 \times 10^{-8}$, one-sided Wilcoxon rank sum). The horizontal line on the box represents the median, box edges show the 25th and 75th percentiles, whiskers represent the minimum and maximum values of the distribution. Throughout the roughly 30s recordings, the extracted latent states converge to being predictive of chronic pain. E) Bar plots of pain state decoding performance (AUC) for all pain metrics in all subjects using LSSM *full models* (all brain regions). One-sided empirical p-values based on permutation tests (n=1000, with multiple comparisons adjustment). *$\mathrm{p}^{\mathrm{FDRadj}} < 0.05$

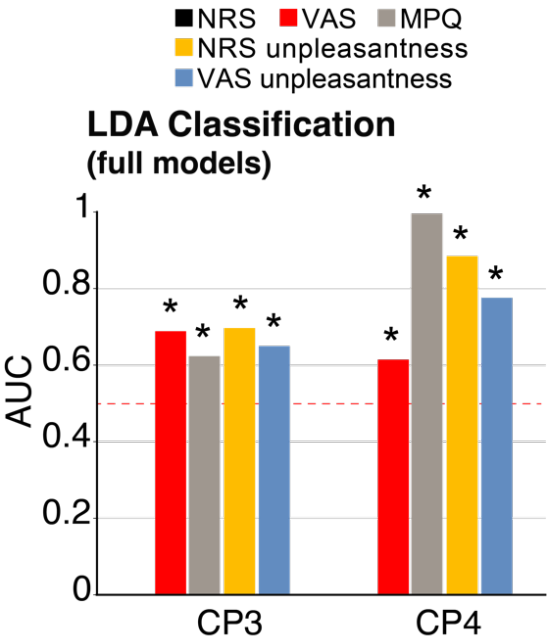


Figure S8. Chronic pain state decoding model performance for alternative pain metrics in two subjects (LDA).
Bar plots of pain state decoding performance (AUC) for alternative pain metrics in two subjects using LDA *full models* (all brain regions). One-sided empirical p-values based on permutation tests (n=1000, without multiple comparisons adjustment), from left to right, are CP3: 0.001, 0.004, 0.003, 0.001; and CP4: 0.002, 0.001, 0.001, 0.001. (see **Figure S4** and **Table S2** for all values)

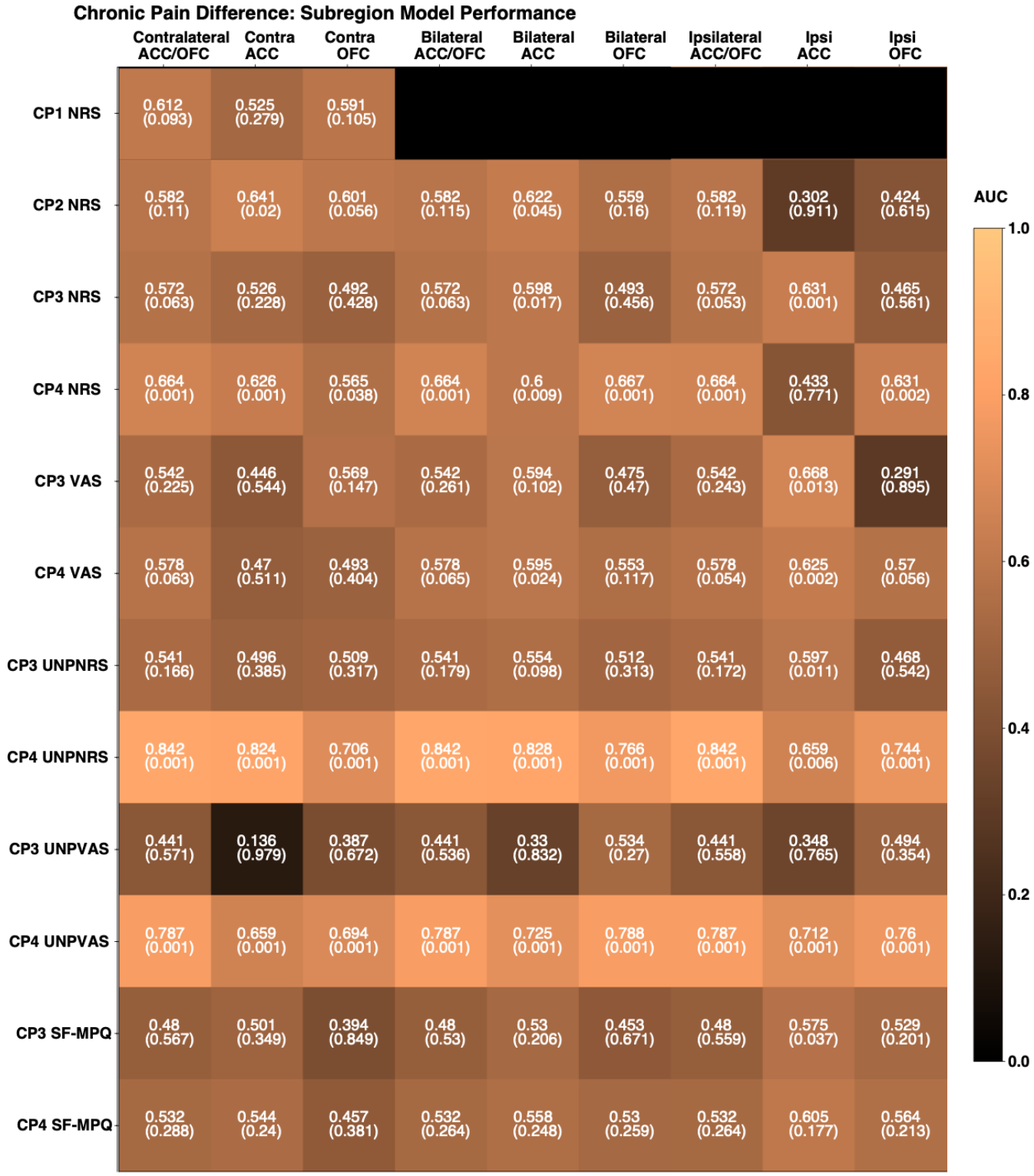


Figure S9a. Chronic Subregion Models for Pain Differences (LDA, Sequential Pain Difference or "Fluctuation")

Heatmap table shows AUC and one-sided empirical p-values based on permutation tests (n=1000, without multiple comparisons adjustment) below in parenthesis for all calculated chronic pain sequential difference models (across columns) and pain score report types (across rows). AUC also represented in color based on color bar. Note this reflects classification of stable pain (no difference) vs increases or decreases (non-zero difference) in successive chronic pain scores. UNP = Unpleasantness

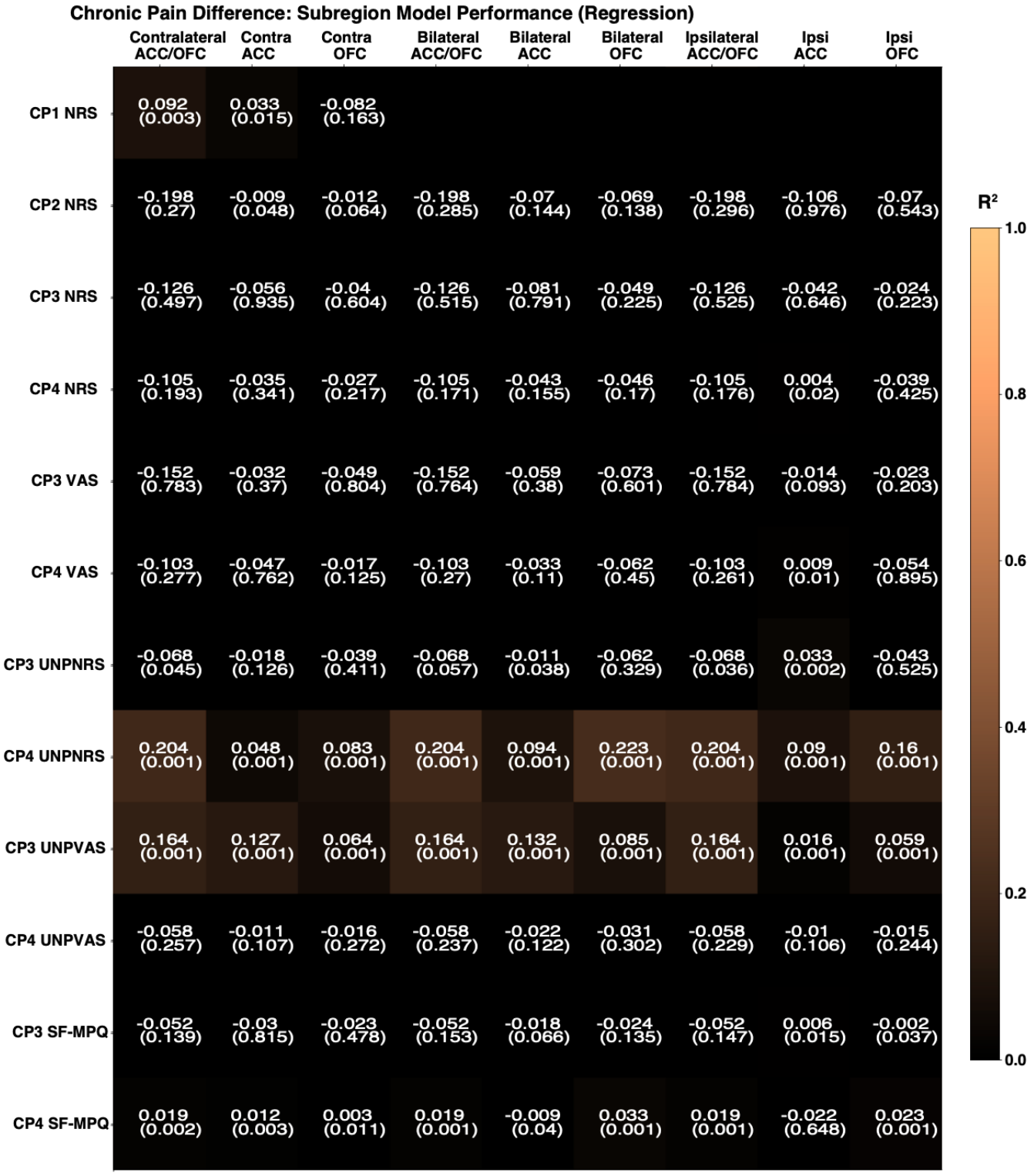


Figure S9b. Chronic Subregion Models for Pain Differences (LASSO, Sequential Pain Difference or "Fluctuation")

Heatmap table shows coefficient of variation (R^2) and one-sided empirical p-values based on permutation tests (n=1000, without multiple comparisons adjustment) below in parenthesis for all calculated chronic pain sequential difference models (across columns) and pain score report types (across rows). R^2 also represented in color based on color bar.

Figure S10. Acute Subregion Model Performance (LDA)

Heatmap table of AUC and one-sided empirical p-values based on permutation tests (n=1000, without multiple comparisons adjustment) below in parenthesis for all calculated models (across columns) and pain score report types (across rows) for acute affected (top panel) and unaffected sides (bottom panel). Red boxes outline significantly decoded pain. AUC represented in color as per color bar on right.

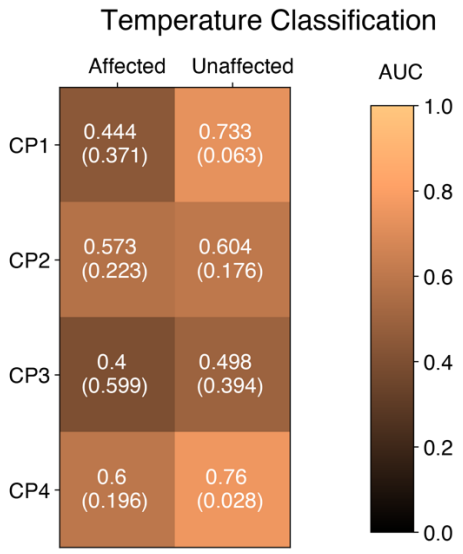


Figure S11. Acute Temperature Classification - Full Model Performance (LDA)

Heatmap table of AUC and one-sided empirical p-values based on permutation tests (n=1000, without multiple comparisons adjustment) below in parenthesis for full models classifying dichotomized thermal probe temperature during Acute Pain when stimulus was applied to affected side and unaffected side. Classification was significant only for CP4 on the unaffected side. AUC value corresponds to color bar on right.

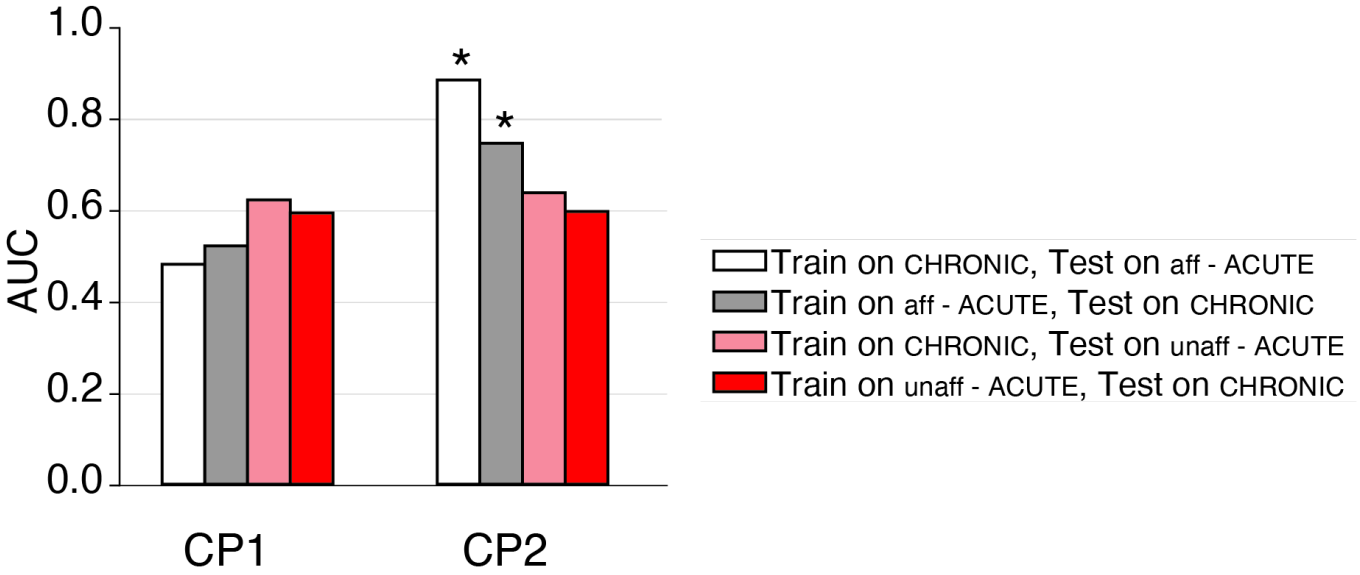


Figure S12. Cross-trained acute pain decoding models show generalization of chronic pain neural features to acute pain states in only one subject (LDA).

Decoder performance for each subject when *full models* trained on acute thermal pain neural features and pain metrics were used to classify chronic pain data. Data for acute pain stimuli applied to the chronic pain affected (aff-ACUTE) side reached significance for one subject, while no significant cross-decoding was seen for unaffected (unaff-ACUTE) side. One-sided empirical p-values based on permutation tests (n=1000, without multiple comparisons adjustment) *p = 0.007

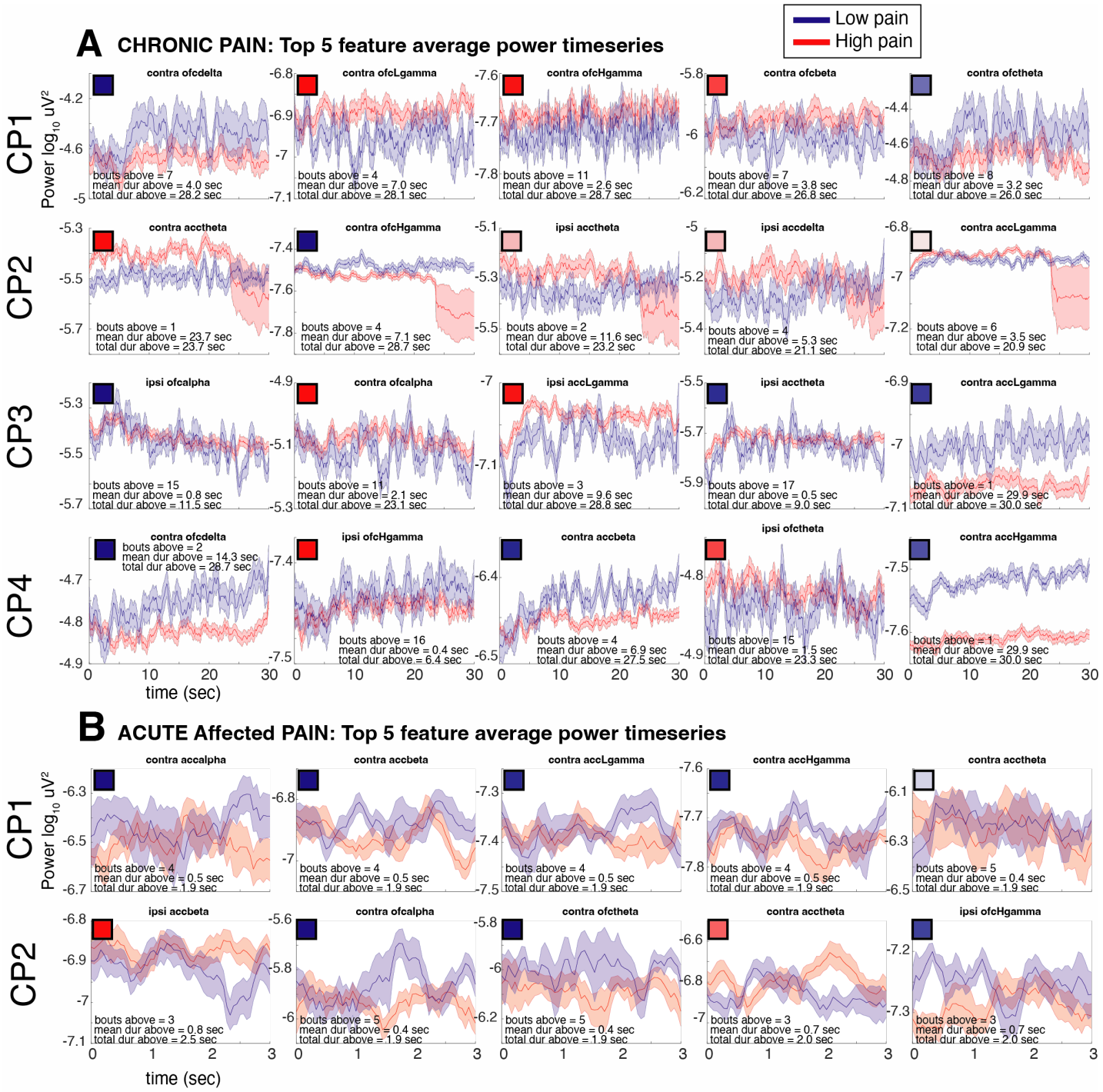


Figure S13. Temporal Dynamics of Top 5 Power Features supporting Acute and Chronic Pain Decoding

Power timeseries plots of top 5 weighted features showing temporal dynamics of relative power increases or decreases during dichotomized high pain states for chronic (A) and acute-affected pain (B) conditions. Red curves show recording clip-averaged curves for high pain reports and blue curves show average for low pain reports (shaded area = s.e.m.). Colored square in top left corner shows feature importance of that feature taken from **Figure 2E** for chronic (A) and **Figure 3G** for acute pain (B). Note that when the feature weight was negatively signed (blue square), average feature power was lower during high pain states and vice versa (red square = positive sign weight). Text at bottom quantifies *bouts* (number of times higher curve crossed above lower curve), *mean bout duration* (mean duration higher curve was above lower curve before passing under), and *total duration above* (total duration higher curve was higher than lower curve). See **Figure 5.**

Diurnal patterns of individual features / powerbands

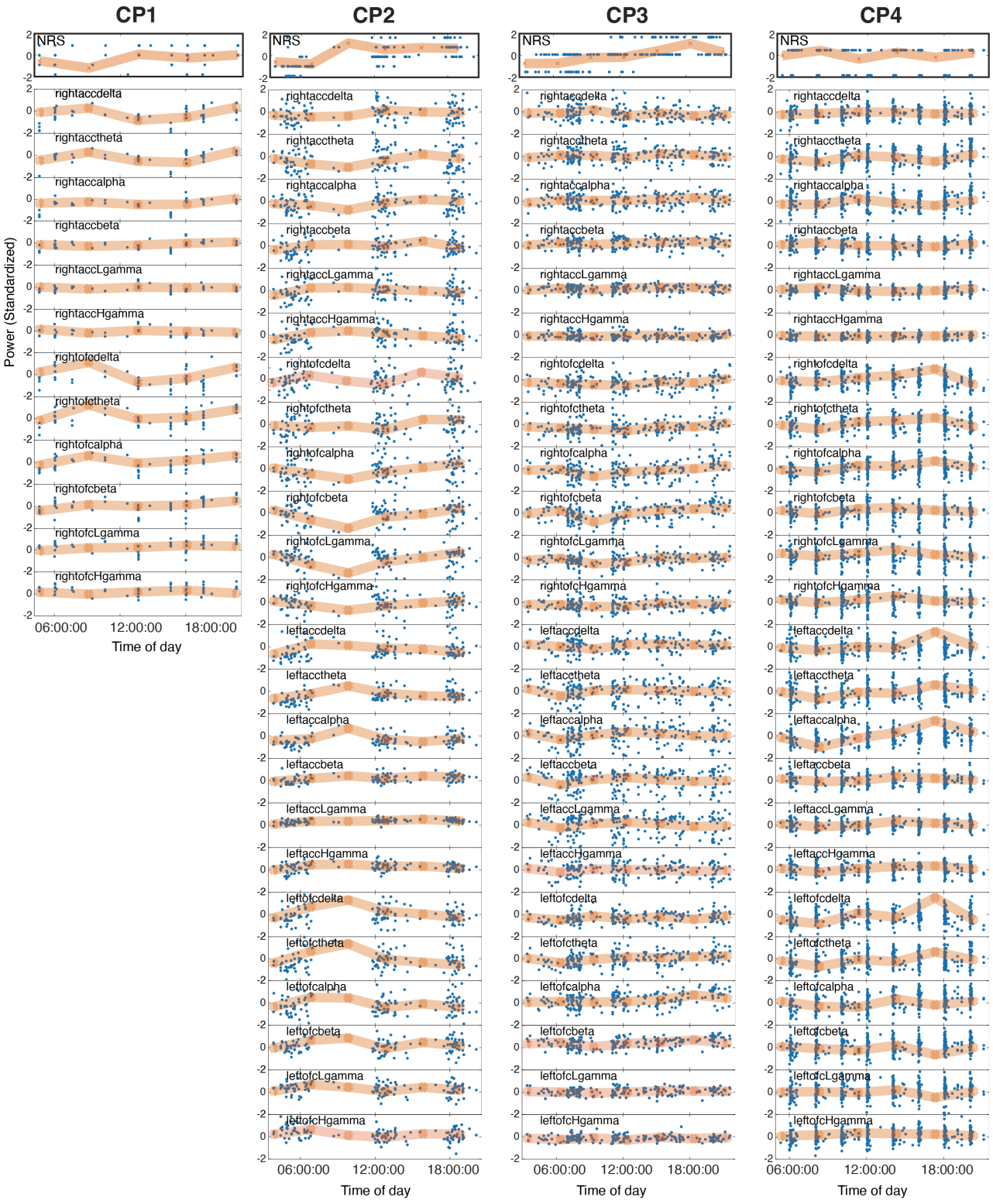


Figure S14. Diurnal patterns of pain intensity NRS and neural power features are not correlated.

Blue points indicate z-scored pain NRS values (top row) or feature power (indicated in top left text) organized by time-of-day of report for each subject (in columns). Red overlying line shows linear trend line of diurnal association, sampled at 3-hour resolution. See **Table S4** for pain NRS trend vs feature trend correlations.

Supplementary Tables

Patient ID		CP1	CP2	CP3	CP4
Demographics	Age (years)	58	63	56	59
	Sex	Female	Female	Male	Male
	Pain Duration (years)	7	4	6	3
	Pain Syndrome	Post-Stroke	Phantom Limb	Post-Stroke	Post-Stroke
	Brain Lesion	Large Right MCA infarct	None	Punctate Left thalamic infarct	Moderate Left PCA and MCA second division infarcts
	Side of Body Affected with Chronic Pain	Left	Right	Right	Right
	Handedness	Right	Right	Right	Right
	Preoperative depression and anxiety (BDI/BAI)	46/40[1]	17/17	4/16	13/7
	MOCA Score	26	30	27	25
Recording Details	Depth lead side/target	Right ACC	Bilateral ACC	Bilateral ACC	Bilateral ACC
	Cortical paddle side/target	Right OFC	Bilateral OFC	Bilateral OFC	Bilateral OFC
	# recordings with pain reports	89	137	234	452
	Pain NRS (mean ± SD)	8.65 ± 1.12	8.29 ± 0.79	6.68 ± 1.06	8.25 ± 0.64
Chronic pain Metrics Reported		pain intensity NRS	pain intensity NRS	pain intensity NRS and VAS, pain unpleasantness NRS and VAS, SF-MPQ	pain intensity NRS and VAS, pain unpleasantness NRS and VAS, SF-MPQ
Active Pain Medications		Acetaminophen, Gabapentin, Oxycodone	none	Duloxetine, Pregabalin	Oxycodone, Pregabalin

Table S1: Patient characteristics, recording/reporting details.

[1] Despite an elevated preoperative score for the BDI/BAI, the subject was recommended for inclusion for the study by the evaluating psychiatrist due to the observation that clinical symptoms of depression and anxiety were adequately managed.

Subject/ Pain Metric	*Threshold*	*Sensitivity*	*Specificity*	*PPV*	*Accuracy (95% CI)*	*AUC (p-value)*
CP1 chronic NRS	0.633	0.79	0.56	0.83	0.729 (0.597, 0.836)	0.596 (0.09)
CP2 chronic NRS	0.699	0.84	0.76	0.87	0.81 (0.734, 0.872)	0.851 (0.001)
CP3 chronic NRS	0.876	0.62	0.77	0.93	0.647 (0.579, 0.71)	0.705 (0.001)
CP4 chronic NRS	0.833	0.69	0.82	0.93	0.717 (0.673, 0.758)	0.802 (0.001)
CP1 acute[2] *(unaffected side)*	0.774	0.77	0.33	0.84	0.688 (0.413, 0.89)	0.385 (0.42)
CP2 acute (unaffected side)	1	0.07	1.0	1.0	0.35 (0.154, 0.592)	0.226 (0.84)
CP3 acute (unaffected side)	0.99	0.21	1.0	1.0	0.5 (0.313, 0.687)	0.545 (0.27)
CP4 acute (unaffected side)	0.061	0.85	0.25	0.55	0.56 (0.349, 0.756)	0.301 (0.80)
CP1 acute (affected side)	0.88	0.43	0.83	0.75	0.615 (0.316, 0.861)	0.476 (0.37)
CP2 acute (affected side)	0.668	0.92	0.88	0.92	0.9 (0.683, 0.988)	0.833 (0.01)
CP3 acute (affected side)	0.981	0.3	0.89	0.83	0.52 (0.313, 0.722)	0.444 (0.51)
CP4 acute (affected side)	0.98	0.15	1.0	1.0	0.56 (0.349, 0.756)	0.365 (0.69)
CP3 chronic unpleasantness NRS	0.915	0.56	0.81	0.93	0.596 (0.527, 0.663)	0.696 (0.003)
CP4 chronic unpleasantness NRS	0.927	0.87	0.85	0.97	0.865 (0.83, 0.895)	0.885 (0.001)
CP3 chronic unpleasantness VAS	0.533	0.65	0.59	0.64	0.62 (0.554, 0.682)	0.649 (0.002)
CP4 chronic unpleasantness VAS	0.855	0.81	0.75	0.94	0.801 (0.757, 0.84)	0.776 (0.001)
CP3 chronic VAS	0.473	0.68	0.63	0.65	0.654 (0.589, 0.715)	0.688 (0.001)
CP4 chronic VAS	0.501	0.81	0.473	0.66	0.661 (0.611, 0.709)	0.615 (0.001)
CP3 chronic SF-MPQ	0.686	0.51	0.75	0.77	0.603 (0.537, 0.666)	0.623 (0.005)
CP4 chronic SF-MPQ	0.956	0.97	0.97	0.99	0.973 (0.951, 0.987)	0.995 (0.001)

Table S2: Decoding Characteristics for all full binary LDA models.

Full models include all available neural data. Threshold indicates optimal LDA decision boundary. PPV = positive predictive value. AUC = area under the curve. CI = Confidence Interval. One-sided P-values were determined based on permutation tests (n=1000, without multiple comparisons adjustment)

[2] Acute pain measured via verbal NRS

	AFFECTED SIDE			UNAFFECTED SIDE		
Subject	Temp (Celsius)	Pain NRS	S.D.	Temp (Celsius)2	Pain NRS	S.D.
CP1	28	5.33	3.79	37	1.00	1.41
	30	7.33	2.08	40	0.00	0.00
	32	8.00	0.00	42	1.40	1.95
	34	8.83	0.29	44	5.00	1.41
	36	8.67	0.58	46	8.50	0.71
CP2	38	0.00	0.00	38	0.50	1.00
	42	1.25	1.50	42	1.00	1.15
	44	1.25	0.96	44	2.25	1.26
	46	5.75	0.96	46	7.00	0.00
	48	7.25	0.50	48	7.00	0.00
CP3	34	0.20	0.45	34	0.00	0.00
	38	1.20	1.10	38	0.20	0.45
	42	2.00	1.87	42	1.50	0.84
	44	3.00	0.00	44	2.20	0.45
	46	8.40	0.55	46	4.0	0.55
CP4	34	1.80	1.30	34	1.60	0.89
	38	4.00	1.22	38	2.20	0.84
	40	3.60	1.14	40	2.20	1.10
	44	7.80	1.64	44	3.40	1.67
	46	10.00	0.00	46	7.40	1.67

Table S3. Acute Pain Calibrated Temperatures and Associated Pain Scores

Table shows temperatures calibrated for each patient for Acute pain thermal task showing mean reported pain NRS score and standard deviation (S.D.). Compare to 'Brief Patient Descriptions' in methods for description of allodynia / hypersensitivity.

Feature	CP1	CP2	CP3	CP4
'rightaccdelta'	-0.48	0.46	-0.44	-0.10
FDR adj p-value	1.00	0.99	0.74	0.94
'rightacctheta'	-0.31	-0.01	0.46	-0.15
	1.00	0.99	0.74	0.93
'rightaccalpha'	0.02	-0.03	0.83	-0.25
	1.00	0.99	0.24	0.89
'rightaccbeta'	0.67	0.43	0.67	0.69
	1.00	0.99	0.57	0.61
'rightaccLgamma'	0.72	0.29	0.64	0.57
	1.00	0.99	0.57	0.61
'rightaccHgamma'	0.37	0.29	0.03	0.15
	1.00	0.99	0.95	0.93
'rightofcdelta'	-0.68	-0.03	0.27	-0.51
	1.00	0.99	0.82	0.61
'rightofctheta'	-0.51	0.33	0.36	-0.65
	1.00	0.99	0.80	0.61
'rightofcalpha'	-0.28	-0.11	0.30	-0.64
	1.00	0.99	0.82	0.61
'rightofcbeta'	0.25	-0.29	0.40	-0.51
	1.00	0.99	0.74	0.61
'rightofcLgamma'	0.38	-0.25	0.40	-0.21
	1.00	0.99	0.74	0.93
'rightofcHgamma'	0.51	-0.36	0.40	-0.09
	1.00	0.99	0.74	0.94
'leftaccdelta'		0.07	0.23	-0.53
		0.99	0.82	0.61
'leftacctheta'		0.53	0.14	-0.64
		0.99	0.90	0.61
'leftaccalpha'		0.49	0.08	-0.43
		0.99	0.90	0.66
'leftaccbeta'		0.80	-0.08	-0.04
		0.99	0.90	0.94
'leftaccLgamma'		0.58	-0.21	-0.06
		0.99	0.82	0.94
'leftaccHgamma'		0.31	-0.23	-0.49
		0.99	0.82	0.61
'leftofcdelta'		0.27	0.11	-0.69
		0.99	0.90	0.61
'leftofctheta'		0.15	0.65	-0.71
		0.99	0.57	0.61
'leftofcalpha'		-0.08	0.91	0.40
		0.99	0.11	0.66
'leftofcbeta'		0.20	0.54	0.59
		0.99	0.72	0.61
'leftofcLgamma'		-0.16	0.59	0.61
		0.99	0.65	0.61
'leftofcHgamma'		-0.66	0.33	-0.41
		0.99	0.80	0.66

Table S4: Lack of correlation between diurnal NRS trend and diurnal feature trends

Pearson's correlation values (in bold) and corresponding p-values (two-sided t-statistic, with correction for multiple comparisons) are shown comparing the diurnal trend of each feature with diurnal trend of pain intensity NRS (**see Figure S14**). There are no significant correlation values.

Supplementary References

1. DeCoster, J., Iselin, A.-M. R. & Gallucci, M. A conceptual and empirical examination of justifications for dichotomization. *Psychological Methods* **14**, 349 (20091207).
2. DeCoster, J., Gallucci, M. & Iselin, A.-M. R. Best Practices for Using Median Splits, Artificial Categorization, and their Continuous Alternatives. *Journal of Experimental Psychopathology* **2**, 197–209 (2011).
3. Farrington, D. P. & Loeber, R. Some benefits of dichotomization in psychiatric and criminological research. *Criminal Behaviour and Mental Health* **10**, 100–122 (2000).