A novel cortical biomarker signature predicts individual pain sensitivity

Nahian S Chowdhury PhD*[1,2],Chuan Bi PhD*[3, 4], Andrew J Furman PhD[5,6],Alan KI Chiang PhD[1,2], Patrick Skippen PhD[1,7,8], Emily Si MSc[1], Samantha K Millard PhD[1,2],Sarah M Margerison BA[5,6], Darrah Spies BSc[5,6], Michael L Keaser BA[5,6] Joyce T Da Silva PhD[5,6], Shuo Chen PhD[3], Siobhan M Schabrun PhD**[9,10], David A Seminowicz PhD**[5,6,11]

*Co-first author

**Co-senior author

[1]Center for Pain IMPACT, Neuroscience Research Australia, Sydney, New South Wales, Australia
[2]University of New South Wales, Sydney, New South Wales, Australia
[3]Division of Biostatistics and Bioinformatics, Department of Epidemiology and Public Health, University of Maryland School of Medicine
[4]Division of Biostatistics, Center for Devices and Radiological Health, U.S. Food and Drug Administration, USA
[5]Department of Neural and Pain Sciences, University of Maryland School of Dentistry, Baltimore, USA
[6]Center to Advance Chronic Pain Research, University of Maryland Baltimore, USA
[7]Data Sciences, Hunter Medical Research Institute, Newcastle, NSW, Australia
[8]School of Medicine and Public Health, University of Newcastle, Newcastle, NSW, Australia
[9]The Gray Centre for Mobility and Activity, Parkwood Institute, St. Joseph's Healthcare London, Canada
[10]School of Physical Therapy, University of Western Ontario, London, Canada
[11]Department of Medical Biophysics, Schulich School of Medicine & Dentistry, University of Western Ontario, London, Canada

Date of Revision: 09/21/2024
Word Count: 2990
Keywords: Pain, Biomarker, Electroencephalography, Transcranial Magnetic Stimulation

Corresponding Author:
Professor David A Seminowicz
Department of Medical Biophysics
Schulich School of Medicine & Dentistry
University of Western Ontario, London, Canada,
E-mail: dseminow@uwo.ca

Key Points

Question: Can individuals be accurately classified as high or low pain sensitive based on two features of cortical activity: sensorimotor peak alpha frequency (PAF) and corticomotor excitability (CME)?

Findings: In a cohort study of 150 healthy participants, the performance of a logistic regression model was outstanding in a training set (n=100) and excellent in a test set (n=50), with the combination of slower PAF and CME depression predicting higher pain. Results were reproduced across a range of methodological parameters, and inclusion of covariates did not improve model performance

Meaning: A novel cortical biomarker comprised of PAF and CME can accurately distinguish high and low pain sensitive individuals

Abstract

Background: Biomarkers would greatly assist chronic pain management. The present study aimed to undertake analytical validation of a sensorimotor cortical biomarker signature for pain consisting of two measures: sensorimotor peak alpha frequency (PAF) and corticomotor excitability (CME), using a human model of prolonged temporomandibular pain (masseter intramuscular injection of nerve growth factor [NGF]).

Methods: 150 participants received an injection of NGF to the right masseter muscle on Days 0 and 2, inducing prolonged pain lasting up to 4 weeks. Electroencephalography (EEG) to assess PAF and transcranial magnetic stimulation (TMS) to assess CME were recorded on Days 0, 2 and 5. We determined the predictive accuracy of the PAF/CME biomarker signature using a nested control-test scheme: machine learning models were run on a training set (n = 100), where PAF and CME were predictors and pain sensitivity was the outcome. The winning classifier was assessed on a test set (n = 50) comparing the predicted pain labels against the true labels.

Results: The winning classifier was logistic regression, with an outstanding area under the curve (AUC=1.00). The locked model assessed on the test set had excellent performance (AUC=0.88). Results were reproduced across a range of methodological parameters. Moreover, inclusion of sex and pain catastrophizing as covariates did not improve model performance, suggesting the model including biomarkers only was more robust. PAF and CME biomarkers showed good-excellent test-retest reliability.

Conclusions: This study provides evidence for a sensorimotor cortical biomarker signature for pain sensitivity. The combination of accuracy, reproducibility, and reliability, suggests the PAF/CME biomarker signature has substantial potential for clinical translation.

Several objective pain biomarkers have been proposed, including neuroimaging markers of mechanistic/structural abnormalities[1-3], neural oscillatory rhythms[4] and "multi-omics" metrics of micro RNA[5], proteins[6] , lipids and metabolites[7]. Such biomarkers would greatly assist decision making in the diagnosis, prevention and treatment of chronic pain[8]. However, attempts at establishing pain biomarkers have suffered from either insufficient sample sizes to conduct full-scale analytical validation using machine learning[8-10], failure to use clinically relevant pain models[11-13] or lack of assessment of reproducibility or test-retest reliability[14,15]. These factors have hindered the clinical translatability of prospective pain biomarkers.

Research suggests that the neural oscillatory rhythms involved in processing nociceptive input, and the corticospinal signalling involved in the subsequent motor response, are both critical in shaping the subjective experience of pain[4,16]. This work has culminated to the identification of a promising sensorimotor cortical biomarker signature for predicting pain sensitivity involving two metrics: 1) sensorimotor peak alpha frequency (PAF), defined as the dominant sensorimotor cortical oscillation in the alpha (8-12Hz) range[17] and 2) corticomotor excitability (CME), defined as the efficacy at which signals are relayed from primary motor cortex (M1) to peripheral muscles[18]. Previous work has shown that slower PAF prior to pain onset and reduced CME during prolonged pain ("depression") are associated with more pain, while faster PAF and increased CME ("facilitation") are associated with less pain[19-23]. Given individuals who experience higher pain in the early stages of a prolonged pain episode (e.g. post-surgery) are more likely to develop chronic pain in the future[24], slow PAF prior to an anticipated prolonged pain episode and/or CME depression during the acute stages of pain could be predictors for the transition to chronic pain.

This paper presents the main outcomes of the PREDICT trial, a pre-registered (NCT04241562[25]) analytical validation of the PAF/CME biomarker signature using a model

of prolonged myofascial temporomandibular pain (masseter intramuscular injection of nerve growth factor [NGF]). Repeated NGF injections induce progressively developing prolonged pain lasting up to 4 weeks[23,26], and has been shown to mimic chronic pain characteristics such as time course (gradual development), type of pain (movement-evoked), functional impairments, hyperalgesia (decreased pressure pain thresholds) and mechanism of sensitization[27,28]. This makes the NGF model a highly standardised prolonged pain model with which to undertake biomarker validation.

The aim of the PREDICT trial was to determine whether individuals could be accurately classified as high or low pain sensitive based on baseline PAF and CME facilitator/depressor classification. We predicted the area under the curve (AUC) of the receiver operator characteristic (ROC) curve for distinguishing high and low pain sensitive individuals would be at least 70% (which represents an acceptable AUC)[29].

Methods

Participants

The PREDICT trial enrolled 159 healthy participants (70 females, 89 males, mean age 25.1 ± 6.1), with 150 participants remaining after dropouts. Ethical approval was obtained from the University of New South Wales (HC190206) and the University of Maryland Baltimore (HP-00085371). Written, informed consent was obtained. The supplementary appendix contains all additional details regarding participant characteristics and methodology.

Experimental Protocol

Outcomes were collected over a period of 30 days. Participants attended the laboratory on Day 0, 2, and 5. Baseline questionnaire data were collected on Day 0. Pressure pain thresholds, PAF and CME were measured on Day 0, 2 and 5. PAF was obtained via a 5-minute eyes-closed resting-state EEG recording from 63 electrodes. Sensorimotor PAF was computed by identifying the component in the signal (transformed by independent component analysis) that had a clear alpha peak (8-12Hz) upon frequency decomposition and a scalp topography suggestive of a sensorimotor source. CME was obtained using transcranial magnetic stimulation (TMS) mapping; single pulses of TMS delivered to the left primary motor cortex (M1), and motor evoked potentials (MEPs) recorded from the right masseter muscle using electromyography (EMG) electrodes. TMS was delivered at each site on a 1cm-spaced grid superimposed over the scalp, and a map of the corticomotor representation of the masseter muscle was generated. Corticomotor excitability was indexed as map volume, which is calculated by summing MEP amplitudes from all "active sites" on the grid. NGF was injected into the right masseter muscle at the end of the Day 0 and 2 laboratory sessions. Electronic pain diaries were collected from Days 1 to 30 at 10am and 7pm each day, where participant rated their pain (0-10) during various activities. Pain upon functional jaw movement is a key criterion for the diagnosis of TMD[30] and pain during chewing and yawning are higher compared to other activities after an NGF injection to the masseter muscle[28,31]. As such, the primary outcomes were pain upon chewing and yawning. The protocol and methodology are shown in Figure 1A and 1B.

Analytical Validation Plan

Division of the Data. Analysis was conducted in R, MATLAB and Python, with code publicly available https://github.com/DrNahianC/PREDICT_Scripts. Figure 1C details the

analysis plan. We adopted a nested-control-testing scheme by partitioning 150 participants into the first 100 (training set) and second 50 (test set) individuals to participate in the study.

Growth Mixture Modelling. We used growth mixture modelling (GMM) in R[32-34] to form two classes: high and low pain sensitive. For this categorization, we used the sum of pain upon chewing and yawning data, and pain data from Days 1-7, as this was the timeframe when pain was most prominent (eFigure 3). As such, participants would more reliably fall into high and low pain sensitive classes during this timeframe. The first and last 40 participants (80 in total) in the training set, based on the ordering of probabilities of the pain intensity trajectory belonging to one of the classes, were labelled as high and low pain sensitive. The trained GMM model, once established, was locked and utilized to label the test set. Consequently 38 out of 50 test set participants were labelled, with a different proportion of high and low pain sensitive (24 high and 14 low pain) compared to the training set since the classifications were based on the probability thresholds established in the training set. These labels were recorded for subsequent comparison with the predicted labels produced by the trained machine learning model.

Machine Learning Model Selection and Fine Tuning. We utilized five machine learning models on the labelled training set —logistic regression, random forest, gradient boosting, support vector machine, and neural network. The dependent variable was pain sensitivity label (high/low) identified from the GMM and independent variables were sensorimotor PAF and ΔCME: the latter was typified as facilitator and depressor, depending on whether they showed an increase or decrease in map volume on Day 5 relative to Day 0, respectively. For each model, we identified optimized parameters through 5-fold cross-validation: we randomly divided the 80 participants into an internal training set of 64 participants (consisting of four equal folds of 16) and a validation set of 16. The optimized models in the internal training set were then employed to predict labels in the validation set to

facilitate model selection. The model with the best performance (area under the curve) on the validation set was then locked in.

Test Set Prediction. The locked machine learning model was assessed on the test set. The participant IDs in this set did not coincide with those in the pain diary data, thereby preserving the double-blind nature of the analysis. By using the ground-truth labels (shuffled), predicted labels (unshuffled), and the shuffling order for the test set, we were able to evaluate the model's performance by comparing the reordered predicted labels against the ground-truth labels established by the GMM. Performance was assessed via receiver operating characteristic (ROC) area under the curve (AUC), with 95% confidence intervals reported. AUC values between 0.7-0.8, 0.8-0.9 and 0.9-1 were considered “acceptable”, “excellent”, and “outstanding” respectively[29].

Results

PAF/CME demonstrated good-excellent test-retest reliability

PAF and ΔCME showed good to excellent test-retest reliability across sessions (eFigures 5 and 7).

Outstanding performance on the training validation set

Figure 2A shows the pain scores for participants in the training and test set classified as high and low pain sensitive. Figure 2B (upper) shows the performances of the models across the internal training and validation sets. Logistic regression was the winning classifier based on its outstanding performance (AUC=1.00[1.00-1.00]) when applied to the validation set (Figure 2B lower), with slower PAF and CME depression predicting higher pain with regression coefficients of -1.25 and -1.27 for PAF and CME respectively.

205 **Excellent performance on the test set**

206 When the locked logistic regression model was applied to the test set, performance
207 (Figure 2C upper) was excellent (AUC=0.88[0.78-0.99]). The optimal probability threshold
208 for being classified as high pain sensitive was 0.40, with an associated sensitivity of 0.875
209 and specificity of 0.79. Applying this 0.40 probability threshold to our data, to be labelled as
210 high pain sensitive, a facilitator would need a PAF<9.56, and a depressor would need a
211 PAF<10.57. Figure 2C (lower) shows the differences in pain scores between participants
212 predicted to have high or low pain. Visually one can observe slower PAF in those predicted
213 to have high vs. low pain sensitivity (Figure 2D), This was statistically significant according
214 to a two-sample t-test ($t(48)=5.8$, $p<.001$). Moreover, one can observe a decrease in CME
215 within the masseter motor maps in those predicted to have high pain (Figure 2E), whereas
216 those predicted to have low pain exhibited an increase in CME. The differences in ΔCME
217 between these groups was statistically significant ($t(48)=2.81$, $p=.007$).

218 **A benefit for a *combined* signature**

219 We reran the models to determine whether the combined PAF/CME signature out-
220 performed each measure individually (eFigure 10). The performance of the PAF-only logistic
221 regression model on the training validation and test set were respectively excellent
222 (AUC=0.95[0.84-1.00]) and outstanding (AUC=0.83[0.70-0.96]). The performance of the
223 CME-only logistic regression model for the training validation and test set were respectively
224 excellent (AUC=0.88[0.69-1.00]) and acceptable (AUC=0.75[0.60-0.91]).

225 **Inclusion of covariates did not improve model performance**

226 We evaluated the performance of the biomarker combined with covariates. As there
227 were many variables, we applied feature selection, i.e. filtering features by inspecting p-
228 values when associating predictors and labels, and using parameter tuning to optimize the

coefficients associated with the filtered features. Five features were subsequently selected and optimized – Sensorimotor PAF, CME, Sex, Pain Catastrophizing Scale (PCS) Total and PCS Helplessness. The associations between labels and biomarkers/covariates in the training vs. test set, and performance of the models are shown in Figure 3A and 3B. When including these five features, the performance of the logistic regression model (regression coefficients of -0.86, -0.69, 0.64, 0.02 and 0.06 for PAF, CME, Sex, PCS Total and PCS Helplessness respectively) was outstanding (AUC=1.00[1.00-1.00]) and excellent (AUC=0.81[0.67-0.95]) for the training validation and test set were respectively. Thus, the model with biomarkers only outperformed the model including covariates.

Results were reproducible across methodological choices

To determine whether our results were robust across different methodological choices, we repeated the analysis using different PAF calculation methods, including component level data (with the sensorimotor component chosen manually or using an automated script) vs. sensor level data (with a sensorimotor region of interest), using different frequency windows (8-12Hz vs. 9-11Hz) and using different CME calculation methods (map volume vs. map area). We found that, regardless of the choices, logistic regression was the best or equal-best performing model when applied to the training-validation set (Figure 3C), with AUCs varying from acceptable (AUC=0.77) to outstanding (AUC=1.00). When the locked models were applied to the test set, performance varied from acceptable (AUC=0.73) to excellent (AUC=0.88) (Figure 3D). Lastly, excellent performance was demonstrated when the data was analysed two other ways (eFigure 11 and 12): where GMM pain labels were established using the whole 30 days rather than the first 7 days (training validation AUC=0.84[0.64-1], test AUC =0.89[0.79-0.99]), and when missing pain diary data was not imputed (training validation AUC=0.81[0.6-1], test AUC=0.89[0.79-0.99]).

Discussion

A full-scale analytical validation of the PAF and CME biomarker signature was conducted using a prolonged pain model. In an initial training set (n=100), we found that a logistic regression was the optimal classifier based on its outstanding performance (AUC=100%), with slower PAF and CME depression predicting higher pain. When this model was applied to an independent test set, the AUC was excellent (AUC=88%). PAF/CME showed good-excellent test-retest reliability, and results were reproduced across a range of methodological parameters. Inclusion of covariates did not improve model performance, suggesting the model including biomarkers only was more robust. Overall, the combination of sample size, pain model validity, and biomarker accuracy, reproducibility and reliability suggest the PAF/CME biomarker signature has substantial potential for clinical translation.

Our results suggest that individuals who have slow PAF prior to an anticipated prolonged pain episode and show corticomotor depression during a prolonged pain episode, are more likely to experience higher pain. Model performance was higher combining the two, suggesting consideration of both ascending sensory and descending motor pain processing mechanisms provides more information regarding pain sensitivity. Note that our study used a cohort of healthy participants with strict inclusion/exclusion criteria and an experimental pain model. While this may limit generalizability to clinical populations, the use of a standardized sample/design is a requirement of pre-clinical analytical validation, and an essential first step in the discovery pipeline toward a clinical biomarker signature. Moreover, there is already evidence that the proposed biomarker is generalizable to clinical contexts. A recent study showed that individuals with slower PAF experienced more pain following a thoracotomy[20]. Furthermore, individuals with lower CME during the acute stages of low back pain were more likely to develop chronic pain at 6-months follow-up[35]. This suggests PAF and CME

shows promise in being used in pre-operative and/or post-operative/post-injury contexts to classify high or low pain sensitive individuals. Given that higher acute pain predicts the development of chronic pain[24] PAF and CME could potentially be used as susceptibility biomarkers for the transition from acute to chronic pain.

There are several aspects of our study which stand out. The first is sample size: with recent advancements in machine learning, it has become possible to conduct analytical validation of pain biomarkers. However, deep learning requires a large amount of labelled samples to conduct rigorous training on validation and test sets[8]. Unfortunately, many pain susceptibility biomarker studies have not been sufficiently sampled to adopt such approaches[9,10], and the ones that have used machine learning failed to reach the sample sizes similar to that of the present study[1,2].

Another strength of our findings is reproducibility. Previous work has shown similar associations between slower PAF and higher upper limb pain, post-operative thoracic pain and chronic pain in various body regions[17,19,20] as well as CME depression and higher upper limb pain, chronic patellofemoral pain and development of chronic low back pain[22,23,36,37]. The present study replicated these results in a model of jaw pain, suggesting the biomarker signature may be generalizable to pain more broadly. Note that some studies have not shown a negative relationship between PAF and pain sensitivity[38,39] or a positive relationship between CME depression and pain sensitivity[31]. However, these studies were not sufficiently sampled to conduct analytical validation of the kind presented in this study. Nonetheless, the mixed findings could also arise from differences in methodological choices in the estimation of PAF e.g. frequency windows[39] and use of sensor vs. component space data[40] and estimation of CME e.g. map volume[22] vs. area[31]. For this reason, we repeated the main analysis using different methodological choices and found at least acceptable AUCs, supporting the reproducibility of our results.

The PAF/CME measures demonstrated good-excellent reliability. Reliability is a desirable characteristic which assists in the widespread application of pain biomarkers[8]. We found that participants exhibit stable PAF across days despite the presence of pain, and even when considering different methodological factors that may influence the reliability such as pre-processing pipeline, recording length and frequency window[14]. Indeed, reliable PAF was found with a recording length as short as 2 minutes and minimal data pre-processing. We also showed that those who show CME depression on Day 2 are also likely to show CME depression on Day 5 (and vice versa for those who show CME facilitation). This was shown even when an automated method of determining MEP amplitude on each trial was applied. Thus, our work not only shows that PAF and CME can predict pain, but the relative ease with which reliable PAF/CME data can be obtained is promising for subsequent clinical translation.

Another strength of this study is our pain model. While other pain biomarker studies have shown promising results, these studies were restricted to pain models utilizing transient painful stimuli lasting seconds to minutes[11-13]. The brief nature of the painful stimuli questions the external validity of these findings and limits generalizability to clinical populations. In contrast, the present study used a prolonged pain model lasting weeks. Several other studies have shown that injections of NGF to the neck, elbow or masseter muscles can mimic symptoms of clinical neck pain[41], chronic lateral epicondylalgia[27] and TMD[28] respectively. Thus, the observed relationships between PAF/CME and pain in the present study show promise in terms of clinical applicability.

Lastly, the PAF/CME biomarker demonstrated high performance. A previous study found that connectivity between medial prefrontal cortex and nucleus accumbens in 39 sub-acute low back pain patients (pain duration 6-12 weeks) could predict future pain persistence at ~7, 29 and 54 weeks, with AUCs of 67-83%[1]. Another study on 24 sub-acute low back

pain patients showed that white matter fractional anisotropy measures in the superior longitudinal fasciculus and internal capsule predicted pain persistence over the next year, with an AUC of 81%[2]. Though the present did not directly assess the transition to chronic pain, our AUCs of 100% (validation set) and 88% (test set) appear comparatively high. We therefore encourage future clinical studies to determine whether PAF/CME can predict the transition from acute to chronic pain.

Conclusions

A novel biomarker signature comprised of PAF and CME accurately and reliably distinguishes high and low pain sensitive individuals during prolonged jaw pain with an excellent AUC of 88% in an independent test set. No other pain biomarker study has shown this combination of biomarker accuracy, reproducibility, reliability and pain model validity, suggesting high potential for clinical translation.

Acknowlegements

This project was funded by the National Institutes of Health (R61 NS113269/NS/NINDS thH HHS/United States). Clinical Trial Registration: NCT04241562 https://classic.clinicaltrials.gov/ct2/show/results/NCT04241562?view=results.
The authors have no conflicts to declare. The corresponding author had full access to all the data in the study and takes responsibility for the integrity of the data and accuracy of the data analysis. The data supporting the findings of this study are available from the corresponding author upon reasonable request.

References

1. Baliki MN, Petre B, Torbey S, et al. Corticostriatal functional connectivity predicts transition to chronic back pain. *Nature neuroscience*. 2012;15(8):1117-1119.
2. Mansour AR, Baliki MN, Huang L, et al. Brain white matter structural properties predict transition to chronic pain. *PAIN®*. 2013;154(10):2160-2168.
3. Vachon-Presseau E, Tétreault P, Petre B, et al. Corticolimbic anatomical characteristics predetermine risk for chronic pain. *Brain*. 2016;139(7):1958-1970.
4. Green A, Wang S, Stein J, et al. Neural signatures in patients with neuropathic pain. *Neurology*. 2009;72(6):569-571.
5. Wang W-T, Zhao Y-N, Han B-W, Hong S-J, Chen Y-Q. Circulating microRNAs identified in a genome-wide serum microRNA expression analysis as noninvasive biomarkers for endometriosis. *The Journal of Clinical Endocrinology & Metabolism*. 2013;98(1):281-289.
6. Singh JA, Noorbaloochi S, Knutson KL. Cytokine and neuropeptide levels are associated with pain relief in patients with chronically painful total knee arthroplasty: a pilot study. *BMC Musculoskeletal Disorders*. 2017;18(1):1-6.
7. Sasamoto N, Zeleznik OA, Vitonis AF, et al. Presurgical blood metabolites and risk of postsurgical pelvic pain in young patients with endometriosis. *Fertility and Sterility*. 2022;117(6):1235-1245.
8. Zhang Z, Gewandter JS, Geha P. Brain imaging biomarkers for chronic pain. *Frontiers in Neurology*. 2022;12:734821.
9. Sluka KA, Wager TD, Sutherland SP, et al. Predicting chronic postsurgical pain: current evidence and a novel program to develop predictive biomarker signatures. *Pain*. 2023;164(9):1912-1926.
10. Eldabe S, Obara I, Panwar C, Caraway D. Biomarkers for chronic pain: significance and summary of recent advances. *Pain Research and Management*. 2022;2022
11. Lee J-J, Kim HJ, Čeko M, et al. A neuroimaging biomarker for sustained experimental and clinical pain. *Nature medicine*. 2021;27(1):174-182.
12. Woo C-W, Schmidt L, Krishnan A, et al. Quantifying cerebral contributions to pain beyond nociception. *Nature communications*. 2017;8(1):14211.
13. Wager TD, Atlas LY, Lindquist MA, Roy M, Woo C-W, Kross E. An fMRI-based neurologic signature of physical pain. *New England Journal of Medicine*. 2013;368(15):1388-1397.
14. Chowdhury NS, Skippen P, Si E, et al. The reliability of two prospective cortical biomarkers for pain: EEG peak alpha frequency and TMS corticomotor excitability. *Journal of Neuroscience Methods*. 2023/02/01/ 2023;385:109766. doi:https://doi.org/10.1016/j.jneumeth.2022.109766
15. Davis KD, Aghaeepour N, Ahn AH, et al. Discovery and validation of biomarkers to aid the development of safe and effective pain therapeutics: challenges and opportunities. *Nature Reviews Neurology*. 2020;16(7):381-400.
16. van Dieen JH, Reeves NP, Kawchuk G, van Dillen LR, Hodges PW. Motor Control Changes in Low Back Pain: Divergence in Presentations and Mechanisms. *The Journal of orthopaedic and sports physical therapy*. Jun 2019;49(6):370-379. doi:10.2519/jospt.2019.7917
17. Furman AJ, Meeker TJ, Rietschel JC, et al. Cerebral peak alpha frequency predicts individual differences in pain sensitivity. *NeuroImage*. 2018;167:203-210.
18. Weavil JC, Amann M. Corticospinal excitability during fatiguing whole body exercise. *Progress in brain research*. 2018;240:219-246.
19. Furman AJ, Prokhorenko M, Keaser ML, et al. Sensorimotor peak alpha frequency is a reliable biomarker of prolonged pain sensitivity. *Cerebral Cortex*. 2020;30(12):6069-6082.
20. Millard SK, Furman AJ, Kerr A, et al. Predicting postoperative pain in lung cancer patients using preoperative peak alpha frequency. *British Journal of Anaesthesia*. 2022;128(6):e346-e348.
21. Seminowicz DA, Thapa T, Furman AJ, et al. Slow peak alpha frequency and corticomotor depression linked to high pain susceptibility in transition to sustained pain. *bioRxiv*. 2018:278598.

22. Seminowicz DA, Thapa T, Schabrun SM. Corticomotor depression is associated with higher pain severity in the transition to sustained pain: A longitudinal exploratory study of individual differences. *The Journal of Pain*. 2019;20(12):1498-1506.
23. Chowdhury NS, Chang W-J, Millard SK, et al. The Effect of Acute and Sustained Pain on Corticomotor Excitability: A Systematic Review and Meta-analysis of group-and individual-level data. *The Journal of Pain*. 2022;23(10):1680-1696.
24. Hah JM, Cramer E, Hilmoe H, et al. Factors associated with acute pain estimation, postoperative pain resolution, opioid cessation, and recovery: secondary analysis of a randomized clinical trial. *JAMA network open*. 2019;2(3):e190168-e190168.
25. Seminowicz DA, Bilska K, Chowdhury NS, et al. A novel cortical biomarker signature for predicting pain sensitivity: protocol for the PREDICT longitudinal analytical validation study. *Pain Reports*. 2020;5(4)
26. Svensson P, Cairns BE, Wang K, Arendt-Nielsen L. Injection of nerve growth factor into human masseter muscle evokes long-lasting mechanical allodynia and hyperalgesia. *Pain*. 2003;104(1-2):241-247.
27. Bergin MJG, Hirata R, Mista C, et al. Movement evoked pain and mechanical hyperalgesia after intramuscular injection of nerve growth factor: a model of sustained elbow pain. *Pain Medicine*. 2015;16(11):2180-2191.
28. Schabrun SM, Si E, Millard SK, et al. Intramuscular injection of nerve growth factor as a model of temporomandibular disorder: nature, time-course, and sex differences characterising the pain experience. *Neurobiology of Pain*. 2023;13:100117.
29. Hosmer Jr DW, Lemeshow S, Sturdivant RX. *Applied logistic regression*. vol 398. John Wiley & Sons; 2013.
30. Schiffman E, Ohrbach R, Truelove E, et al. Diagnostic criteria for temporomandibular disorders (DC/TMD) for clinical and research applications: recommendations of the International RDC/TMD Consortium Network and Orofacial Pain Special Interest Group. *Journal of oral & facial pain and headache*. 2014;28(1):6.
31. Costa YM, Exposto FG, Kothari M, et al. Masseter corticomotor excitability is decreased after intramuscular administration of nerve growth factor. *Eur J Pain*. Oct 2019;23(9):1619-1630. doi:10.1002/ejp.1438
32. Proust-Lima C, Philipps V, Liquet B. Estimation of extended mixed models using latent classes and latent processes: the R package lcmm. *arXiv preprint arXiv:150300890*. 2015;
33. Li F, Duncan TE, Duncan SC, Acock A. Latent growth modeling of longitudinal data: A finite growth mixture modeling approach. *Structural Equation Modeling*. 2001;8(4):493-530.
34. Ram N, Grimm KJ. Methods and measures: Growth mixture modeling: A method for identifying differences in longitudinal change among unobserved groups. *International journal of behavioral development*. 2009;33(6):565-576.
35. Jenkins LC, Chang W-J, Buscemi V, et al. Cortical function and sensorimotor plasticity are prognostic factors associated with future low back pain after an acute episode: the UPWaRD prospective cohort study. *Pain*. 2022:10.1097.
36. Jenkins LC, Chang W-J, Buscemi V, et al. Low Somatosensory Cortex Excitability in the Acute Stage of Low Back Pain Causes Chronic Pain. *The Journal of Pain*. 2022/02/01/ 2022;23(2):289-304. doi:https://doi.org/10.1016/j.jpain.2021.08.003
37. Te M, Baptista AF, Chipchase LS, Schabrun SM. Primary motor cortex organization is altered in persistent patellofemoral pain. *Pain Medicine*. 2017;18(11):2224-2234.
38. Nir R-R, Sinai A, Raz E, Sprecher E, Yarnitsky D. Pain assessment by continuous EEG: association between subjective perception of tonic pain and peak frequency of alpha oscillations during stimulation and at rest. *Brain research*. 2010;1344:77-86.
39. De Martino E, Gregoret L, Zandalasini M, Graven-Nielsen T. Slowing in peak-alpha frequency recorded after experimentally-induced muscle pain is not significantly different between high and low pain-sensitive subjects. *The Journal of Pain*. 2021;22(12):1722-1732.

40. Millard SK, Chiang AKI, Humburg P, et al. Effects of nicotine compared to placebo gum on sensitivity to pain and mediating effects of peak alpha frequency. eLife Sciences Publications, Ltd; 2023.

41. Christensen SWM, Elgueta-Cancino E, Simonsen MB, et al. Effect of prolonged experimental neck pain on exercise-induced hypoalgesia. *Pain*. 2022;163(12):2411-2420.

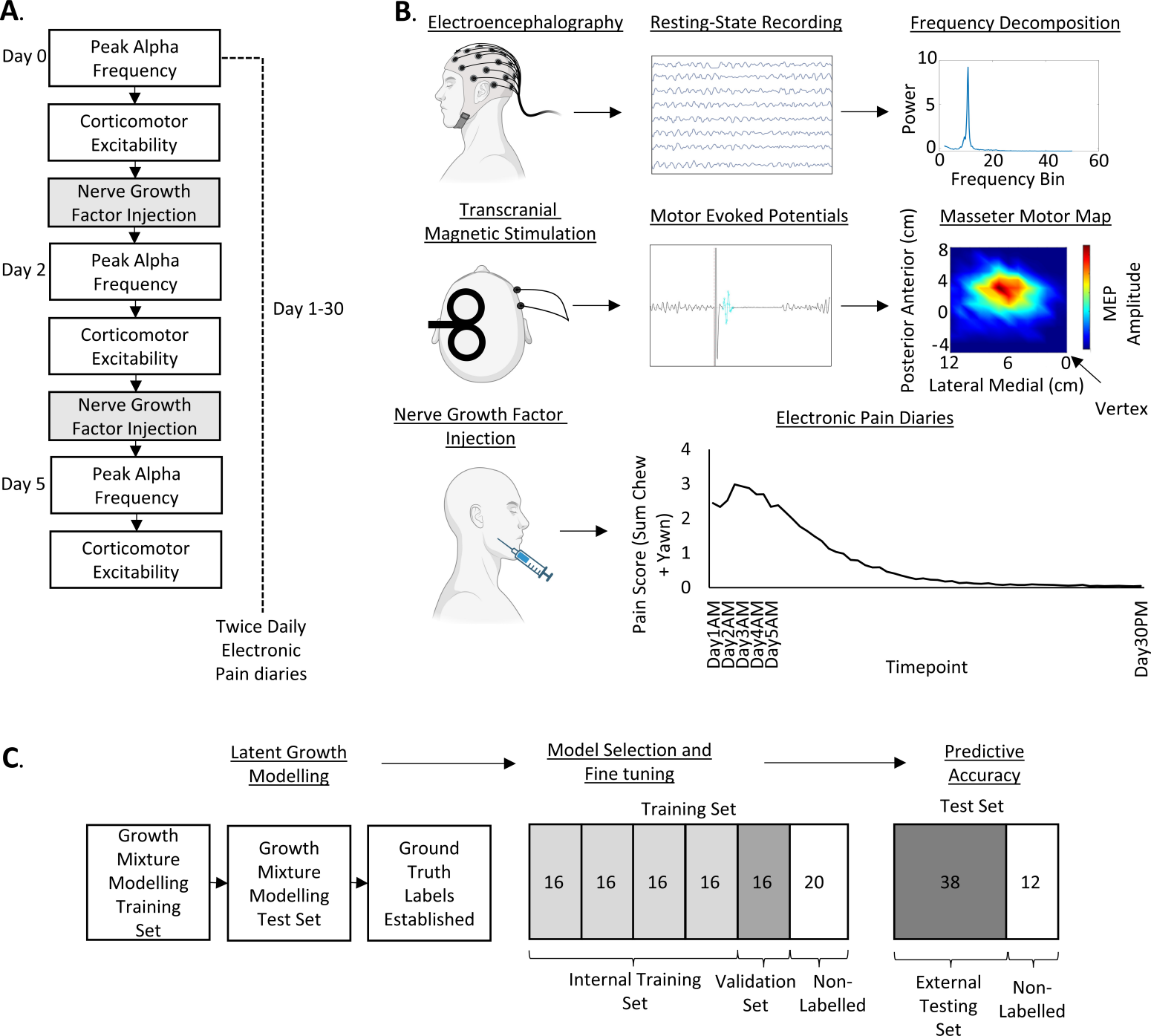
Figure Captions

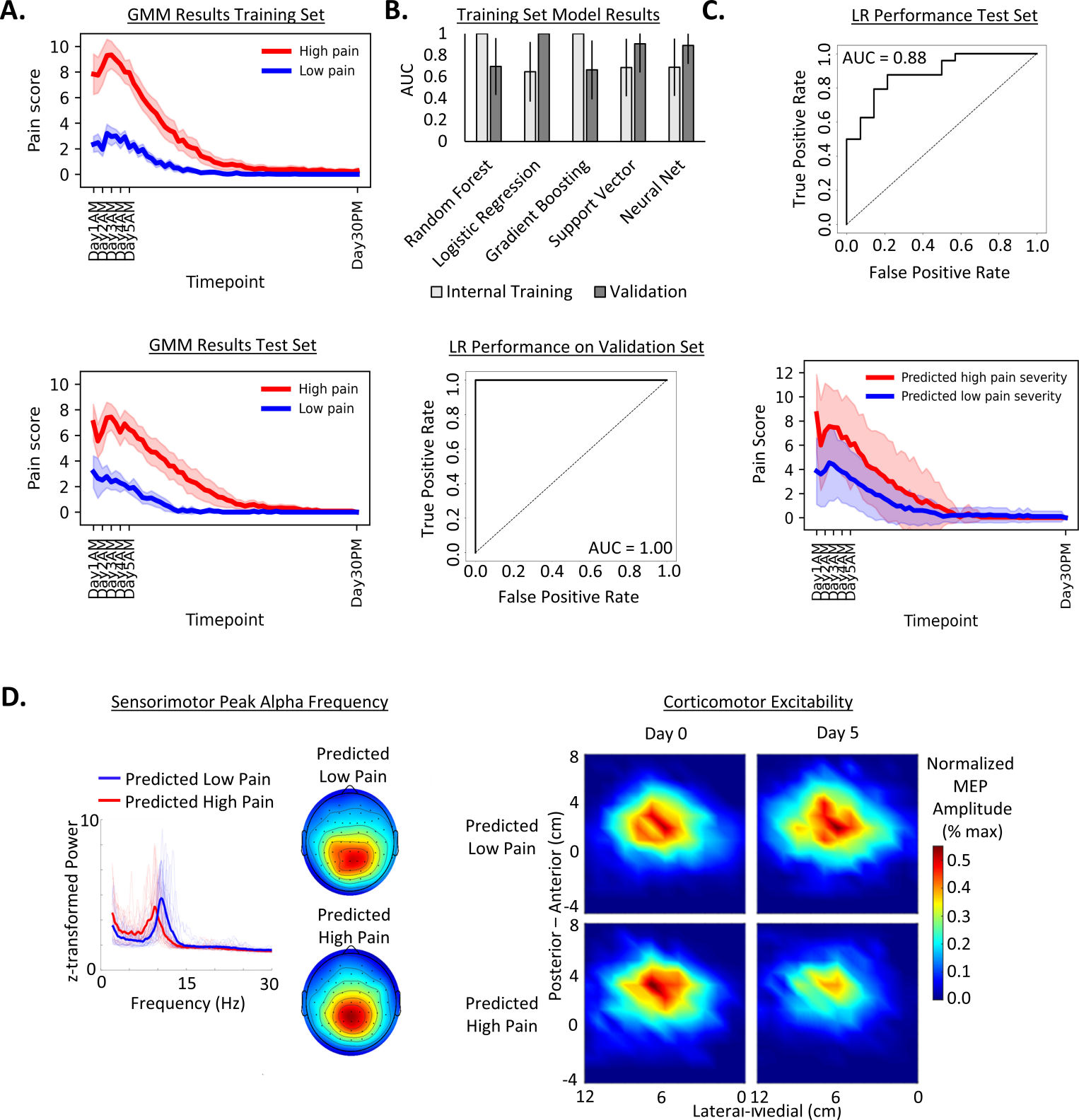
Figure 1. **Study methodology, including study design, data collection procedures and analysis plan. (A)** Experimental protocol showing timeline of data collection procedures. On Day 0, we measured peak alpha frequency (PAF) and corticomotor excitability (CME). At the end of the session, an injection of nerve growth factor (NGF) was administered to the right masseter muscle. On Day 2, PAF and CME were measured, followed by a second NGF injection. On Day 5, PAF and CME were measured. From Days 1-30, electronic diaries measuring jaw pain were sent to participants at 10AM and 7PM each day. **(B)** Details of the methodology. Sensorimotor PAF was measured using a 5 minutes eyes closed resting state EEG recording. Sensorimotor PAF was computed by identifying the component in the signal (transformed by independent component analysis) that had a clear alpha peak in the 8–12 Hz range upon frequency decomposition and a scalp topography suggestive of a source predominately over the sensorimotor cortex. TMS mapping was conducted by stimulating the scalp area over left M1 to obtain a map of the representation of the right masseter muscle. The map consists of the motor-evoked potential (MEP) amplitude at each stimulated location, with CME corresponding to the map volume (sum of all MEPs from active sites). **(C)** Details of the analysis plan. We adopted a nested-control-test scheme by partitioning the 150 subjects into a training set consisting of 100 subjects and an independent test set of 50 subjects. We labelled a subset of participants in the training (n = 80) and test set (n = 38) as high or low pain sensitive using growth mixture modelling (GMM) to establish "ground-truth" labels. We then ran various machine learning models on the labelled training set (with PAF/CME as predictors, and pain severity labels as outcome), and determined optimized parameters through 5-fold cross-validation i.e. randomly dividing the 80 subjects into an internal training set of 64 subjects (with 4 equal folds of 16) and a validation set of 16. The optimized models in the internal training set were employed to predict labels in the validation set to facilitate model selection. The model with the best performance on the validation set was then locked in, and applied to the labelled test set, comparing the predicted labels of high/low pain sensitive with the ground-truth labels of high/low pain sensitive.

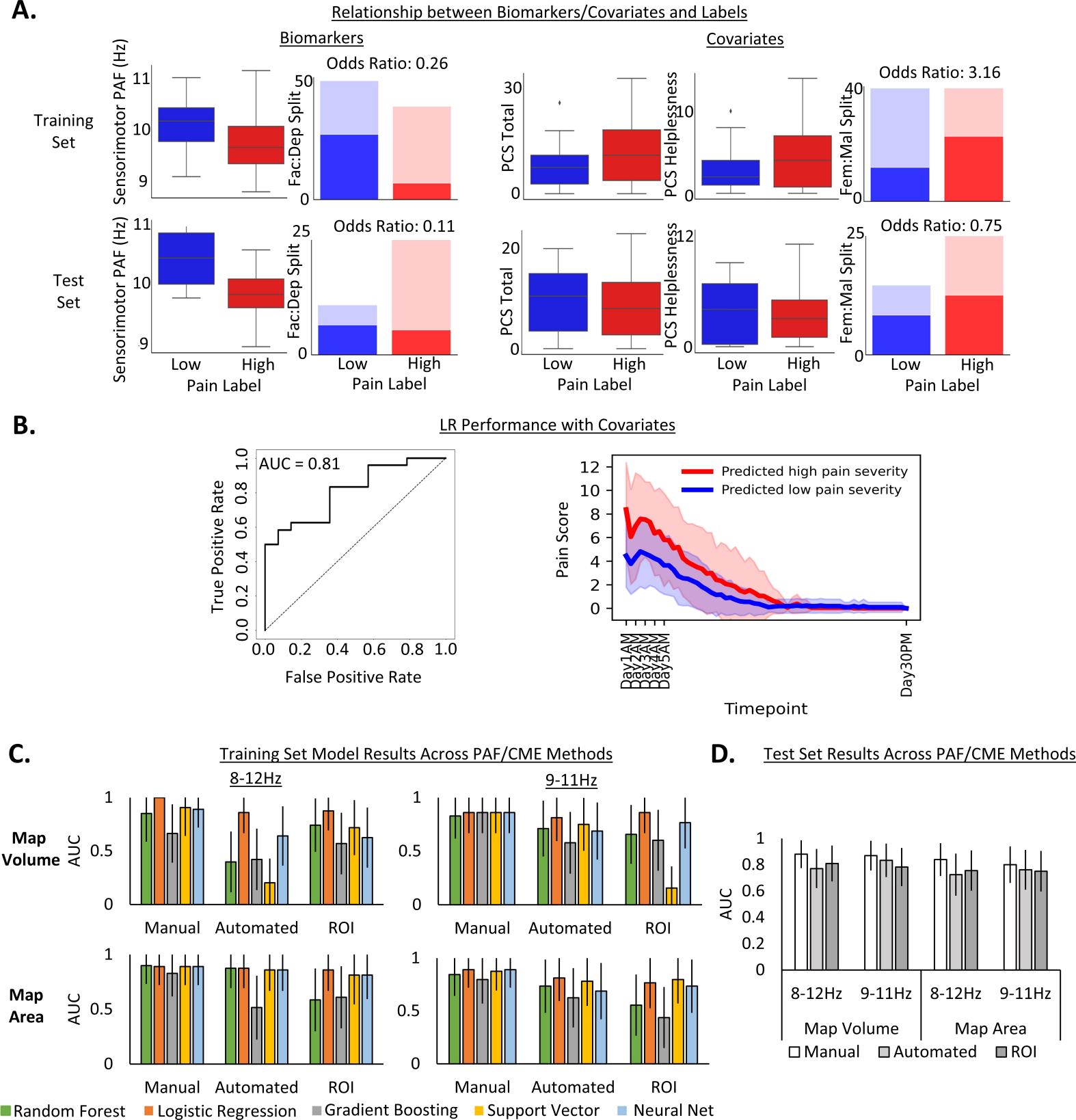
Figure 2. **Performance of the combined PAF and CME biomarker on training and test set. (A)** Results of the growth mixture modelling which categorized 80 participants in the training set (left) and 38 participants in the test (right) as high or low pain sensitive. Data shows mean pain score (chew + yawn pain rating) for each timepoint, while the shaded area shows 95% confidence intervals. **(B)** The upper panel shows performances (AUC [95% confidence intervals]) of various machine learning models for the internal training set and validation set. Logistic regression (LR) was chosen as the optimal classifier based on outstanding AUC of 100% as shown in the lower panel. **(C)** The upper panel shows the performance of the locked logistic regression model when applied to the test set, which was in the excellent range (AUC of 88%). The lower panel shows the pain trajectories (mean chew + yawn pain and 95% confidence intervals) of participants predicted to have high or low pain sensitivity based on the locked logistic regression model. **(D)** Individual and mean z-transformed spectral plots and topography of the sensorimotor alpha component on Day 0 for participants predicted to have high or pain sensitivity based on the locked logistic regression model. **(E)** The mean motor cortex maps on Day 0 and Day 5 showing normalized motor evoked potential (MEP) amplitude (expressed as a proportion of the maximal MEP

amplitude) for participants predicted to have high or low pain sensitivity based on the locked logistic regression model.

Figure 3. **Performance of the combined PAF and CME biomarker on training and test set when including covariates, and across PAF/CME calculation methods. (A)** Visualisation of biomarkers and covariates for the training and test sets across high (red) and low (blue) pain labels identified from the GMM. Data on PAF, PCS total and PCS helplessness are plotted as boxplots, while data on CME and Sex are plotted according to facilitator: depressor (Fac: dep) and female: male (fem: mal) split respectively, including odd ratios. A lower odds ratio means a lower probability of high pain sensitive individuals belonging to the facilitator or female categories. For PAF and CME, low pain was associated with fast PAF and CME facilitation for both training and test sets. In contrast, the relationship between covariates and labels were in the opposite direction for the training and test set, suggesting the relationship between biomarkers and labels was consistent. **(B)** The left panel shows the performance of the locked logistic regression model on the test set when including covariates in the model. The right panel shows pain trajectories (mean chew + yawn score and 95% confidence intervals) of participants predicted to have high or low pain sensitivity based on the locked logistic regression model including covariates. **(C)** The performance of each machine learning model (AUC [95% confidence intervals]) on the training validation set across different PAF/CME calculation methods. This includes the sensorimotor component chosen manually after an independent component analysis, component identified using an automated script after an independent component analysis, or using a sensorimotor region of interest (ROI, mean of Cz, C3 and C4) in electrode space, to calculate PAF. We also looked at different frequency windows for computing PAF (8-12Hz vs. 9-11Hz) or CME calculated using map area or map volume. **(D)** The performance of the locked logistic regression model (AUC [95% confidence intervals]) when applied to the test set, across different PAF/CME calculation methods.


A.
Day 0
Peak Alpha Frequency
Corticomotor Excitability
Nerve Growth Factor Injection
Day 2
Peak Alpha Frequency
Corticomotor Excitability
Nerve Growth Factor Injection
Day 5
Peak Alpha Frequency
Corticomotor Excitability
Day 1-30
Twice Daily Electronic Pain diaries
B.
Electroencephalography
Resting-State Recording
Frequency Decomposition
Power
Frequency Bin
Transcranial Magnetic Stimulation
Motor Evoked Potentials
Masseter Motor Map
Posterior Anterior (cm)
Lateral Medial (cm)
MEP Amplitude
Vertex
Nerve Growth Factor Injection
Electronic Pain Diaries
Pain Score (Sum Chew + Yawn)
Day1AM
Day2AM
Day3AM
Day4AM
Day5AM
Day30PM
Timepoint
C.
Latent Growth Modelling
Model Selection and Fine tuning
Predictive Accuracy
Growth Mixture Modelling Training Set
Growth Mixture Modelling Test Set
Ground Truth Labels Established
Training Set
16
16
16
16
16
20
Internal Training Set
Validation Set
Non-Labelled
Test Set
38
12
External Testing Set
Non-Labelled


A.
GMM Results Training Set
High pain
Low pain
Pain score
Day1AM
Day2AM
Day3AM
Day4AM
Day5AM
Day30PM
Timepoint
B.
Training Set Model Results
AUC
Random Forest
Logistic Regression
Gradient Boosting
Support Vector
Neural Net
Internal Training
Validation
C.
LR Performance Test Set
AUC = 0.88
True Positive Rate
False Positive Rate
GMM Results Test Set
LR Performance on Validation Set
AUC = 1.00
Predicted high pain severity
Predicted low pain severity
Pain Score
D.
Sensorimotor Peak Alpha Frequency
Predicted Low Pain
Predicted High Pain
z-transformed Power
Frequency (Hz)
Corticomotor Excitability
Day 0
Day 5
Normalized MEP Amplitude (% max)
Posterior – Anterior (cm)
Lateral-Medial (cm)


A.
Relationship between Biomarkers/Covariates and Labels
Biomarkers
Covariates
Training Set
Test Set
Sensorimotor PAF (Hz)
Fac:Dep Split
PCS Total
PCS Helplessness
Fem:Mal Split
Odds Ratio: 0.26
Odds Ratio: 3.16
Odds Ratio: 0.11
Odds Ratio: 0.75
Low
High
Pain Label
B.
LR Performance with Covariates
AUC = 0.81
True Positive Rate
False Positive Rate
Predicted high pain severity
Predicted low pain severity
Pain Score
Day1AM
Day2AM
Day3AM
Day4AM
Day5AM
Day30PM
Timepoint
C.
Training Set Model Results Across PAF/CME Methods
8-12Hz
9-11Hz
Map Volume
Map Area
AUC
Manual
Automated
ROI
Random Forest
Logistic Regression
Gradient Boosting
Support Vector
Neural Net
D.
Test Set Results Across PAF/CME Methods
8-12Hz
9-11Hz
Map Volume
Map Area
Manual
Automated
ROI