

1 **A novel cortical biomarker signature accurately and reliably predicts individual pain**
2 **sensitivity: The PREDICT longitudinal analytical validation study**

3 Nahian S Chowdhury*^{1,2}, Chuan Bi*^{3,4}, Andrew J Furman^{5,6}, Alan KI Chiang^{1,2}, Patrick
4 Skippen^{1,7,8}, Emily Si¹, Samantha K Millard^{1,2}, Sarah M Margerison^{5,6}, Darrah Spies^{5,6},
5 Michael L Keaser^{5,6}, Joyce T Da Silva^{5,6}, Shuo Chen³, Siobhan M Schabrun**^{1,9,10}, David A
6 Seminowicz**^{1,5,6,11}

7 *Co-first author

8 **Co-senior author

9 ¹Center for Pain IMPACT, Neuroscience Research Australia, Sydney, New South Wales,
10 Australia

11 ²University of New South Wales, Sydney, New South Wales, Australia

12 ³Division of Biostatistics and Bioinformatics, Department of Epidemiology and Public
13 Health, University of Maryland School of Medicine

14 ⁴Division of Biostatistics, Center for Devices and Radiological Health, U.S. Food and Drug
15 Administration, USA

16 ⁵Department of Neural and Pain Sciences, University of Maryland School of Dentistry,
17 Baltimore, USA

18 ⁶Center to Advance Chronic Pain Research, University of Maryland Baltimore, USA

19 ⁷Data Sciences, Hunter Medical Research Institute, Newcastle, NSW, Australia

20 ⁸School of Medicine and Public Health, University of Newcastle, Newcastle, NSW, Australia

21 ⁹The Gray Centre for Mobility and Activity, Parkwood Institute, London, Canada

22 ¹⁰School of Physical Therapy, University of Western Ontario, London, Canada

23 ¹¹Department of Medical Biophysics, Schulich School of Medicine & Dentistry, University
24 of Western Ontario, London, Canada

25

26

27 **Data availability statement:** The data supporting the findings of this study are available
28 from the corresponding author upon reasonable request

29 **Keywords:** Pain, Biomarker, Electroencephalography, Transcranial Magnetic Stimulation

30 **Disclosures:** This project was funded by the National Institutes of Health (R61
31 NS113269/NS/NINDS NIH HHS/United States). The authors have no conflicts to declare.

32 **Corresponding Author:**

33 Professor David A Seminowicz

34 Department of Medical Biophysics

35 Schulich School of Medicine & Dentistry

36 University of Western Ontario, London, Canada,

37 E-mail: dseminow@uwo.ca

38

39

40

41

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69

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 and inclusion of covariates in the modelling. PAF and CME biomarkers showed good-excellent test-retest reliability.

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

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

79 Recent evidence shows promise for a sensorimotor cortical biomarker signature for
80 predicting the severity of a prolonged pain episode. The biomarker signature reflects
81 individual differences in ascending sensory and descending motor processing, comprising
82 two metrics: 1) sensorimotor peak alpha frequency (PAF), defined as the dominant
83 sensorimotor cortical oscillation in the alpha (8-12Hz) range [16], and is related to the
84 efficiency in which the brain can inhibit incoming sensory input [17, 18], and 2) corticomotor
85 excitability (CME), defined as the efficacy at which signals are relayed from primary motor
86 cortex (M1) to peripheral muscles [19]. CME is altered during pain as individuals adopt
87 different movement strategies to cope with pain [20, 21]. Previous work has shown that
88 slower PAF prior to pain onset and reduced CME during prolonged pain (“depression”) are
89 associated with more pain, while faster PAF and increased CME (“facilitation”) are
90 associated with less pain [21-25]). Given individuals who experience higher pain in the early
91 stages of a prolonged pain episode (e.g. post-surgery) are more likely to develop chronic pain
92 in the future [26], slow PAF prior to an anticipated prolonged pain episode and/or CME
93 depression during the acute stages of pain are potential predictors for the transition to chronic
94 pain.

95 This paper presents the main outcomes of the PREDICT trial, a pre-registered
96 (NCT04241562, [27]) full-scale analytical validation of the PAF/CME biomarker signature
97 using a human model of prolonged myofascial temporomandibular pain (masseter
98 intramuscular injection of nerve growth factor [NGF]). Repeated NGF injections induce
99 progressively developing prolonged pain lasting up to 4 weeks [25, 28], and has been shown
100 to mimic chronic pain characteristics such as time course (gradual development), type of pain
101 (movement-evoked), functional impairments, hyperalgesia (increased pressure pain
102 thresholds) and mechanism of sensitization [29, 30]. This makes the NGF model a highly

103 standardised and clinically relevant prolonged pain model with which to undertake biomarker
104 validation.

105 The aim of the PREDICT trial was to determine whether individuals could be
106 accurately classified as high or low pain sensitive based on baseline PAF and early CME
107 facilitator/depressor classification. We predicted the area under the curve (AUC) of the
108 receiver operator characteristic (ROC) curve for distinguishing high and low pain sensitive
109 individuals would be at least 70% (which represents an acceptable AUC) [31].

110

111 **Methods**

112 **Participants**

113 The PREDICT trial enrolled 159 healthy participants (70 females, 89 males, mean age
114 25.1 ± 6.1), with 150 participants remaining after participant dropouts. Ethical approval was
115 obtained from the University of New South Wales (HC190206) and the University of
116 Maryland Baltimore (HP-00085371). Written, informed consent was obtained. The
117 supplementary appendix contains all additional details regarding participant characteristics
118 and methodology.

119 **Experimental Protocol**

120 Outcomes were collected over a period of 30 days. Participants attended the laboratory
121 on Day 0, 2, and 5. Baseline questionnaire data were collected on Day 0. Pressure pain
122 thresholds, PAF and CME were measured on Day 0, 2 and 5. PAF was obtained via a 5-minute
123 eyes-closed resting-state EEG recording from 63 scalp electrodes. Sensorimotor PAF was
124 computed by identifying the component in the signal (transformed by independent component
125 analysis) that had a clear alpha peak (8-12Hz) upon frequency decomposition and a scalp
126 topography suggestive of a sensorimotor source. CME was obtained using transcranial
127 magnetic stimulation (TMS) mapping; single pulses of TMS delivered to the left primary motor
128 cortex (M1), and motor evoked potentials (MEPs) recorded from the right masseter muscle
129 using electromyography (EMG) electrodes. TMS was delivered at each site on a 1cm-spaced
130 grid superimposed over the scalp, and a map of the corticomotor representation of the masseter
131 muscle was generated. Corticomotor excitability was indexed as map volume, which is
132 calculated by summing MEP amplitudes from all “active sites” on the grid. NGF was injected
133 into the right masseter muscle at the end of the Day 0 and 2 laboratory sessions. Electronic pain

134 diaries were collected from Days 1 to 30 at 10am and 7pm each day, where participant rated
135 their pain (0-10) during various activities. Pain upon functional jaw movement is a key criterion
136 for the diagnosis of TMD [32]. Moreover, previous research has shown that, after an NGF
137 injection to the masseter muscle, pain during chewing and yawning are higher compared to
138 other activities [30, 33]. As such, the primary outcomes were pain upon chewing and yawning.
139 The protocol and methodology are shown in Figure 1A and 1B.

140 **Analytical Validation Plan**

141 **Division of the Data.** Analysis was conducted in R, MATLAB and Python, with code
142 publicly available https://github.com/DrNahianC/PREDICT_Scripts. Figure 1C details the
143 analysis plan. We adopted a nested-control-testing scheme by partitioning 150 participants into
144 a training and test set of 100 and 50 participants respectively.

145 **Growth Mixture Modelling.** We used growth mixture modelling (GMM) in R [34-36]
146 to form two participant classes: high and low pain sensitive. For this categorization, we used
147 the sum of pain upon chewing and yawning data, and pain diary trajectories from Days 1-7 for
148 the classification, as this was the timeframe when pain was most prominent (Supplementary
149 Figure 3). As such, participants would more reliably fall into high and low pain sensitive classes
150 during this timeframe. The first and last 40 participants (80 in total) in the training set, based
151 on the ordering of probabilities of the pain intensity trajectory belonging to one of the classes,
152 were then labelled as high and low pain sensitive. The trained GMM model, once established,
153 was locked and utilized to label the test set. Consequently, 38 out of 50 test set participants (24
154 high and 14 low pain) were labelled. These labels were recorded for subsequent comparison
155 with the predicted labels produced by the trained machine learning model.

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

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

168 **Test Set Prediction.** The locked machine learning model was assessed on the test set.
169 The participant IDs in this set did not coincide with those in the pain diary data, thereby
170 preserving the double-blind nature of the analysis. By using the ground truth labels
171 (shuffled), predicted labels (unshuffled), and the shuffling order for the test set, we were able
172 to evaluate the model's performance by comparing the reordered predicted labels against the
173 ground truth labels established by the GMM. Performance was assessed via receiver
174 operating characteristic (ROC) area under the curve (AUC), with 95% confidence intervals
175 reported. AUC values between 0.7-0.8, 0.8-0.9 and 0.9-1 were considered “acceptable”,
176 “excellent”, and “outstanding” respectively [31].

177

178 **Results**

179 **PAF/CME demonstrated good-excellent test-retest reliability**

180 PAF and Δ CME showed good to excellent test-retest reliability across sessions
181 (Supplementary Figures 5 and 7).

182 **Outstanding performance on the training validation set**

183 Figure 2A shows the pain scores for participants in the training and test set classified
184 as high and low pain sensitive based on GMM. Figure 2B (upper) shows the performances of
185 the machine learning models across the internal training and validation sets. Logistic
186 regression was chosen as the optimal classifier based on its outstanding performance
187 (AUC=1.00[1.00-1.00]) when applied to the validation set (Figure 2B lower), with slower
188 PAF and CME depression predicting higher pain.

189 **Excellent performance on the test set**

190 When the locked logistic regression model was applied to the test set, performance
191 (Figure 2C upper) was excellent (AUC=0.88[0.78-0.99]). Figure 2C (lower) shows the
192 differences in pain scores between participants predicted to have high or low pain. Visually
193 one can observe slower peak alpha frequency in those predicted to have high vs. low pain
194 sensitivity (Figure 2D), which was further confirmed with a two-sample t-test ($t(48)=5.8$,
195 $p<.001$). Moreover, one can observe a decrease in CME within the masseter motor maps in

196 those predicted to have high pain (Figure 2E), whereas those predicted to have low pain
197 exhibit an increase in CME. The differences in the change in CME relative to Day 0 between
198 these groups was further confirmed with a two-sample t-test ($t(48)=2.81, p=.007$).

199 **A benefit for a *combined* signature**

200 We reran the models to determine whether the combined PAF/CME signature out-
201 performed each measure individually (Supplementary Figure 9). For PAF alone, the
202 performance of the logistic regression model on the training validation and test set were
203 respectively excellent (AUC=0.95[0.84-1.00]) and outstanding (AUC=0.83[0.70-0.96]). For
204 CME alone, the performance of the logistic regression model for the training validation and
205 test set were respectively excellent (AUC=0.88[0.69-1.00]) and acceptable (AUC=0.75[0.60-
206 0.91]).

207 **Results were reproducible when including covariates**

208 We evaluated the performance of the biomarker combined with demographic and
209 clinical attributes. As we collected a large amount of this data, we applied feature selection,
210 i.e. filtering features by inspecting p-values when associating predictors and labels, and using
211 parameter tuning to optimize the coefficients associated with the filtered features. Five
212 features were subsequently selected and optimized – Sensorimotor PAF, CME, Sex, Pain
213 Catastrophizing Scale (PCS) Total and PCS Helplessness. The associations between labels
214 and biomarkers/covariates in the training vs. test set, and performance of the models are
215 shown in Figure 3A and 3B. When including these five features, the performance of the
216 logistic regression model was for the training validation and test set were respectively
217 outstanding (AUC=1.00[1.00-1.00]) and excellent (AUC=0.81[0.67-0.95]).

218 **Results were reproducible across methodological choices**

219 To determine whether our results were robust across different methodological
220 choices, we repeated the analysis using PAF calculated using component level data (with the
221 sensorimotor component chosen manually or using an automated script) vs. sensor level data
222 (with a sensorimotor region of interest), using different frequency windows (8-12Hz vs. 9-
223 11Hz) and using different CME calculation methods (map volume vs. map area). We found
224 that, regardless of the choices, logistic regression was the best or equal-best performing
225 model when applied to the validation set (Figure 3C), with AUCs varying from acceptable
226 (0.77) to outstanding (1.00). When the locked models were applied to the test set,

227 performance varied from acceptable (AUC=0.73) to excellent (AUC=0.88) (Figure 3D).
228 Lastly, excellent performance was demonstrated when the data was analysed two other ways
229 (Supplementary Figure 10 and 11): where GMM pain labels were established using the whole
230 30 days rather than the first 7 days (training validation AUC=0.84[0.64-1], test AUC
231 =0.89[0.79-0.99]), and when missing pain diary data was not imputed (training validation
232 AUC=0.81[0.6-1], test AUC=0.89[0.79-0.99]).

233

234

Discussion

235 A full-scale analytical validation of the PAF and CME biomarker signature was
236 conducted using a clinically relevant prolonged pain model. In an initial training set (n=100),
237 we found that a logistic regression was the optimal classifier based on its outstanding
238 performance (AUC=100%), with slower PAF and CME depression predicting higher pain.
239 When this model was applied to an independent test set, the AUC was excellent (88%).
240 PAF/CME showed good-excellent test-retest reliability, and results were reproduced across a
241 range of methodological parameters and consideration of covariates. Overall, the combination
242 of sample size, pain model validity, and biomarker accuracy, reproducibility and reliability
243 suggest the PAF/CME biomarker signature has substantial potential for clinical translation.

244 Our results suggest that individuals who have slow PAF prior to an anticipated
245 prolonged pain episode and show corticomotor depression during a prolonged pain episode,
246 are more likely to experience higher pain. Model performance was higher combining the two,
247 suggesting consideration of both ascending sensory and descending motor pain processing
248 mechanisms provides more information regarding pain sensitivity. Overall, we believe this
249 biomarker could be particularly useful in contexts such as predicting post-operative pain. For
250 example, a recent study showed that individuals with slower PAF experienced more pain
251 following a thoracotomy [25]. Given that higher acute pain post-surgery predicts the
252 development of chronic pain [26], our findings suggest individuals with slow PAF/reduced
253 CME could be more likely transition to chronic pain. Indeed, individuals who show lower
254 CME during the acute stages of low back pain were more likely to develop chronic pain at 6-
255 months follow-up [37]. These preliminary findings, along with our analytical validation
256 study, suggest PAF and CME could be susceptibility biomarkers for the transition from acute
257 to chronic pain.

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

265 Another strength of our findings is reproducibility. The majority of work has shown
266 similar associations between higher pain and slower PAF [16, 22] and CME depression [21,
267 25] in models of upper limb pain. The present study replicated these results in a model of
268 prolonged jaw pain, suggesting these associations hold across pain locations. It is important
269 to note that some studies have not shown a negative relationship between PAF and pain
270 sensitivity [38, 39] or a positive relationship between CME depression and pain sensitivity
271 [33]. However, these studies were not sufficiently sampled to conduct analytical validation of
272 the kind presented in this study. Nonetheless, the mixed findings could also arise from
273 differences in methodological choices in the estimation of PAF e.g. frequency windows [39]
274 and use of sensor vs. component space data [40] and estimation of CME e.g. map volume
275 [21] vs. area [33]. For this reason, we repeated the main analysis using different
276 methodological choices and found at least acceptable AUCs. In addition, we found that the
277 inclusion of covariates such as pain catastrophizing and sex did not alter our results, further
278 supporting the reproducibility of our results.

279 The PAF/CME measures demonstrated good-excellent reliability. Reliability is a
280 highly desirable characteristic which assists in the widespread application of pain biomarkers
281 [8]. We found in the present analysis, and previously [14], that participants exhibit stable
282 PAF across days despite the presence of pain, and even when considering different
283 methodological factors that may influence the reliability of PAF such as pre-processing
284 pipeline, recording length and frequency window. Indeed, we found reliable PAF with a
285 recording length as short as 2 minutes and minimal data pre-processing. We also showed that
286 those who show CME depression on Day 2 are also likely to show CME depression on Day 5
287 (and vice versa for those who show CME facilitation). This finding was shown even when an
288 automated method of determining MEP amplitude on each trial was applied. Thus, our work
289 not only shows that PAF and CME can predict pain, but the relative ease with which reliable
290 PAF/CME data can be obtained is promising for subsequent clinical translation.

291 Another strength of this study is the clinical relevance of our pain model, making
292 clinical translation of the current findings highly feasible. While other pain biomarker studies
293 have shown promising results, these studies were restricted to pain models utilizing transient
294 painful stimuli lasting seconds to minutes [11-13]. The brief nature of the painful stimuli
295 questions the external validity of these findings and limits generalizability to clinical
296 populations. In contrast, the present study used a prolonged pain model lasting weeks.
297 Several other studies have shown that injections of NGF to the neck, elbow or masseter
298 muscles can mimic symptoms of clinical neck pain [41], chronic lateral epicondylalgia [29]
299 and TMD [30] respectively. Thus, the observed relationships between PAF/CME and pain in
300 the present study show promise in terms of clinical applicability.

301 Lastly, the PAF/CME biomarker demonstrated high performance. A previous study
302 found that connectivity between medial prefrontal cortex and nucleus accumbens in 39 sub-
303 acute low back pain patients (pain duration 6-12 weeks) could predict future pain persistence
304 at ~7, 29 and 54 weeks, with AUCs of 67-83% [1]. Another study on 24 sub-acute low back
305 pain patients showed that white matter fractional anisotropy measures in the superior
306 longitudinal fasciculus and internal capsule predicted pain persistence over the next year,
307 with an AUC of 81% [2]. Though the present did not directly assess the transition to chronic
308 pain, our AUCs of 100% (validation set) and 88% (test set) appear comparatively high. We
309 therefore encourage future clinical studies to determine whether PAF/CME can predict the
310 transition from acute to chronic pain.

311 **Conclusions**

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

317

318

319

320

321

References

- 322 1. Baliki, M.N., et al., *Corticostriatal functional connectivity predicts transition to chronic back*
323 *pain*. Nature neuroscience, 2012. **15**(8): p. 1117-1119.
- 324 2. Mansour, A.R., et al., *Brain white matter structural properties predict transition to chronic*
325 *pain*. PAIN®, 2013. **154**(10): p. 2160-2168.
- 326 3. Musto, A.E., et al., *Hippocampal neuro-networks and dendritic spine perturbations in*
327 *epileptogenesis are attenuated by neuroprotectin d1*. PLoS One, 2015. **10**(1): p. e0116543.
- 328 4. Vachon-Presseau, E., et al., *Corticolimbic anatomical characteristics predetermine risk for*
329 *chronic pain*. Brain, 2016. **139**(7): p. 1958-1970.
- 330 5. Wang, W.-T., et al., *Circulating microRNAs identified in a genome-wide serum microRNA*
331 *expression analysis as noninvasive biomarkers for endometriosis*. The Journal of Clinical
332 Endocrinology & Metabolism, 2013. **98**(1): p. 281-289.
- 333 6. Singh, J.A., S. Noorbaloochi, and K.L. Knutson, *Cytokine and neuropeptide levels are*
334 *associated with pain relief in patients with chronically painful total knee arthroplasty: a pilot*
335 *study*. BMC Musculoskeletal Disorders, 2017. **18**(1): p. 1-6.
- 336 7. Sasamoto, N., et al., *Presurgical blood metabolites and risk of postsurgical pelvic pain in*
337 *young patients with endometriosis*. Fertility and Sterility, 2022. **117**(6): p. 1235-1245.
- 338 8. Zhang, Z., J.S. Gewandter, and P. Geha, *Brain imaging biomarkers for chronic pain*. Frontiers
339 in Neurology, 2022. **12**: p. 734821.
- 340 9. Sluka, K.A., et al., *Predicting chronic postsurgical pain: current evidence and a novel program*
341 *to develop predictive biomarker signatures*. Pain, 2023. **164**(9): p. 1912-1926.
- 342 10. Eldabe, S., et al., *Biomarkers for chronic pain: significance and summary of recent advances*.
343 Pain Research and Management, 2022. **2022**.
- 344 11. Lee, J.-J., et al., *A neuroimaging biomarker for sustained experimental and clinical pain*.
345 Nature medicine, 2021. **27**(1): p. 174-182.
- 346 12. Woo, C.-W., et al., *Quantifying cerebral contributions to pain beyond nociception*. Nature
347 communications, 2017. **8**(1): p. 14211.
- 348 13. Wager, T.D., et al., *An fMRI-based neurologic signature of physical pain*. New England
349 Journal of Medicine, 2013. **368**(15): p. 1388-1397.
- 350 14. Chowdhury, N.S., et al., *The reliability of two prospective cortical biomarkers for pain: EEG*
351 *peak alpha frequency and TMS corticomotor excitability*. Journal of Neuroscience Methods,
352 2023. **385**: p. 109766.
- 353 15. Davis, K.D., et al., *Discovery and validation of biomarkers to aid the development of safe and*
354 *effective pain therapeutics: challenges and opportunities*. Nature Reviews Neurology, 2020.
355 **16**(7): p. 381-400.
- 356 16. Furman, A.J., et al., *Cerebral peak alpha frequency predicts individual differences in pain*
357 *sensitivity*. NeuroImage, 2018. **167**: p. 203-210.
- 358 17. Mazaheri, A., D.A. Seminowicz, and A.J. Furman, *Peak alpha frequency as a candidate*
359 *biomarker of pain sensitivity: the importance of distinguishing slow from slowing*.
360 NeuroImage, 2022. **262**: p. 119560.
- 361 18. Van Diepen, R.M., J.J. Foxe, and A. Mazaheri, *The functional role of alpha-band activity in*
362 *attentional processing: the current zeitgeist and future outlook*. Current opinion in
363 psychology, 2019. **29**: p. 229-238.
- 364 19. Weavil, J.C. and M. Amann, *Corticospinal excitability during fatiguing whole body exercise*.
365 Progress in brain research, 2018. **240**: p. 219-246.
- 366 20. Summers, S.J., et al., *Motor adaptation varies between individuals in the transition to*
367 *sustained pain*. Pain, 2019. **160**(9): p. 2115-2125.
- 368 21. Seminowicz, D.A., T. Thapa, and S.M. Schabrun, *Corticomotor depression is associated with*
369 *higher pain severity in the transition to sustained pain: A longitudinal exploratory study of*
370 *individual differences*. The Journal of Pain, 2019. **20**(12): p. 1498-1506.

- 371 22. Furman, A.J., et al., *Sensorimotor peak alpha frequency is a reliable biomarker of prolonged*
372 *pain sensitivity*. Cerebral Cortex, 2020. **30**(12): p. 6069-6082.
- 373 23. Millard, S.K., et al., *Predicting postoperative pain in lung cancer patients using preoperative*
374 *peak alpha frequency*. British Journal of Anaesthesia, 2022. **128**(6): p. e346-e348.
- 375 24. Seminowicz, D.A., et al., *Slow peak alpha frequency and corticomotor depression linked to*
376 *high pain susceptibility in transition to sustained pain*. bioRxiv, 2018: p. 278598.
- 377 25. Chowdhury, N.S., et al., *The Effect of Acute and Sustained Pain on Corticomotor Excitability:*
378 *A Systematic Review and Meta-analysis of group-and individual-level data*. The Journal of
379 Pain, 2022. **23**(10): p. 1680-1696.
- 380 26. Hah, J.M., et al., *Factors associated with acute pain estimation, postoperative pain*
381 *resolution, opioid cessation, and recovery: secondary analysis of a randomized clinical trial*.
382 JAMA network open, 2019. **2**(3): p. e190168-e190168.
- 383 27. Seminowicz, D.A., et al., *A novel cortical biomarker signature for predicting pain sensitivity:*
384 *protocol for the PREDICT longitudinal analytical validation study*. Pain Reports, 2020. **5**(4).
- 385 28. Svensson, P., et al., *Injection of nerve growth factor into human masseter muscle evokes*
386 *long-lasting mechanical allodynia and hyperalgesia*. Pain, 2003. **104**(1-2): p. 241-247.
- 387 29. Bergin, M.J.G., et al., *Movement evoked pain and mechanical hyperalgesia after*
388 *intramuscular injection of nerve growth factor: a model of sustained elbow pain*. Pain
389 Medicine, 2015. **16**(11): p. 2180-2191.
- 390 30. Schabrun, S.M., et al., *Intramuscular injection of nerve growth factor as a model of*
391 *temporomandibular disorder: nature, time-course, and sex differences characterising the*
392 *pain experience*. Neurobiology of Pain, 2023. **13**: p. 100117.
- 393 31. Hosmer Jr, D.W., S. Lemeshow, and R.X. Sturdivant, *Applied logistic regression*. Vol. 398.
394 2013: John Wiley & Sons.
- 395 32. Schiffman, E., et al., *Diagnostic criteria for temporomandibular disorders (DC/TMD) for*
396 *clinical and research applications: recommendations of the International RDC/TMD*
397 *Consortium Network and Orofacial Pain Special Interest Group*. Journal of oral & facial pain
398 and headache, 2014. **28**(1): p. 6.
- 399 33. Costa, Y.M., et al., *Masseter corticomotor excitability is decreased after intramuscular*
400 *administration of nerve growth factor*. Eur J Pain, 2019. **23**(9): p. 1619-1630.
- 401 34. Proust-Lima, C., V. Philipps, and B. Liqueur, *Estimation of extended mixed models using latent*
402 *classes and latent processes: the R package lcmm*. arXiv preprint arXiv:1503.00890, 2015.
- 403 35. Li, F., et al., *Latent growth modeling of longitudinal data: A finite growth mixture modeling*
404 *approach*. Structural Equation Modeling, 2001. **8**(4): p. 493-530.
- 405 36. Ram, N. and K.J. Grimm, *Methods and measures: Growth mixture modeling: A method for*
406 *identifying differences in longitudinal change among unobserved groups*. International
407 journal of behavioral development, 2009. **33**(6): p. 565-576.
- 408 37. Jenkins, L.C., et al., *Cortical function and sensorimotor plasticity are prognostic factors*
409 *associated with future low back pain after an acute episode: the UPWARD prospective cohort*
410 *study*. Pain, 2022: p. 10.1097.
- 411 38. Nir, R.-R., et al., *Pain assessment by continuous EEG: association between subjective*
412 *perception of tonic pain and peak frequency of alpha oscillations during stimulation and at*
413 *rest*. Brain research, 2010. **1344**: p. 77-86.
- 414 39. De Martino, E., et al., *Slowing in peak-alpha frequency recorded after experimentally-induced*
415 *muscle pain is not significantly different between high and low pain-sensitive subjects*. The
416 Journal of Pain, 2021. **22**(12): p. 1722-1732.
- 417 40. Millard, S.K., et al., *Effects of nicotine compared to placebo gum on sensitivity to pain and*
418 *mediating effects of peak alpha frequency*. 2023, eLife Sciences Publications, Ltd.
- 419 41. Christensen, S.W.M., et al., *Effect of prolonged experimental neck pain on exercise-induced*
420 *hypoaesthesia*. Pain, 2022. **163**(12): p. 2411-2420.

421

422

Figure Captions

423 Figure 1. **(A)** Experimental protocol showing timeline of data collection procedures. On Day
424 0, we measured peak alpha frequency (PAF) and corticomotor excitability (CME). At the end
425 of the session, an injection of nerve growth factor (NGF) was administered to the right
426 masseter muscle. On Day 2, PAF and CME were measured, followed by a second NGF
427 injection. On Day 5, PAF and CME were measured. From Days 1-30, electronic diaries
428 measuring jaw pain were sent to participants at 10AM and 7PM each day. **(B)** Details of the
429 methodology. Sensorimotor PAF was measured using a 5 minutes eyes closed resting state
430 EEG recording. Sensorimotor PAF was computed by identifying the component in the signal
431 (transformed by independent component analysis) that had a clear alpha peak in the 8–12 Hz
432 range upon frequency decomposition and a scalp topography suggestive of a source
433 predominately over the sensorimotor cortex. TMS mapping was conducted by stimulating the
434 scalp area over left M1 to obtain a map of the representation of the right masseter muscle.
435 The map consists of the motor-evoked potential (MEP) amplitude at each stimulated location,
436 with CME corresponding to the map volume (sum of all MEPs from active sites). **(C)** Details
437 of the analysis plan. We adopted a nested-control-test scheme by partitioning the 150 subjects
438 into a training set consisting of 100 subjects and an independent test set of 50 subjects. We
439 labelled a subset of participants in the training ($n = 80$) and test set ($n = 38$) as high or low
440 pain sensitive using growth mixture modelling (GMM) to establish “ground-truth” labels. We
441 then ran various machine learning models on the labelled training set (with PAF/CME as
442 predictors, and pain severity labels as outcome), and determined optimized parameters
443 through 5-fold cross-validation i.e. randomly dividing the 80 subjects into an internal training
444 set of 64 subjects (with 4 equal folds of 16) and a validation set of 16. The optimized models
445 in the internal training set were employed to predict labels in the validation set to facilitate
446 model selection. The model with the best performance on the validation set was then locked
447 in, and applied to the labelled test set, comparing the predicted labels of high/low pain
448 sensitive with the ground-truth labels of high/low pain sensitive.

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

466

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

489

490

491

492

493

494

495

496

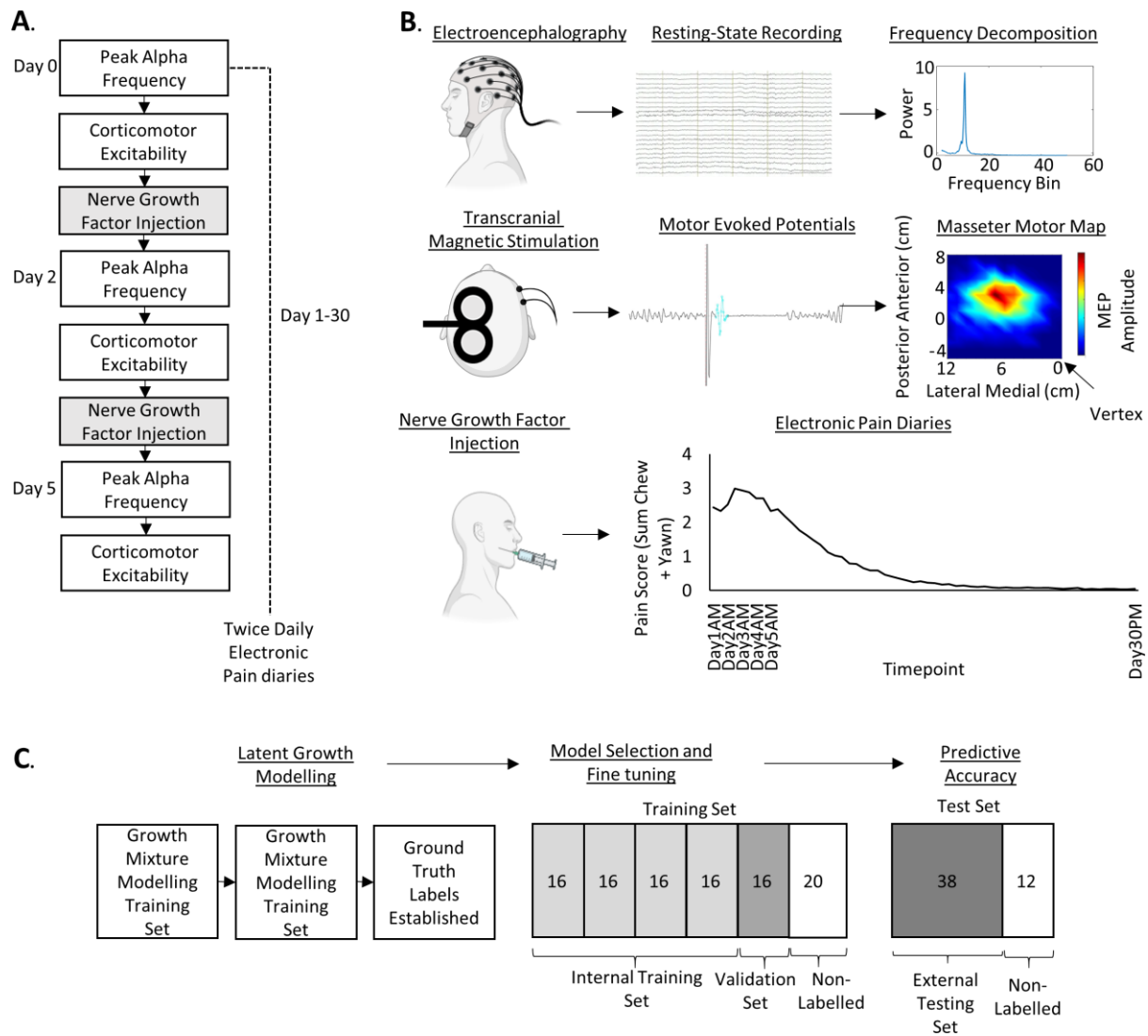
497

498

499

Figures

500 **Figure 1**



501

502

503

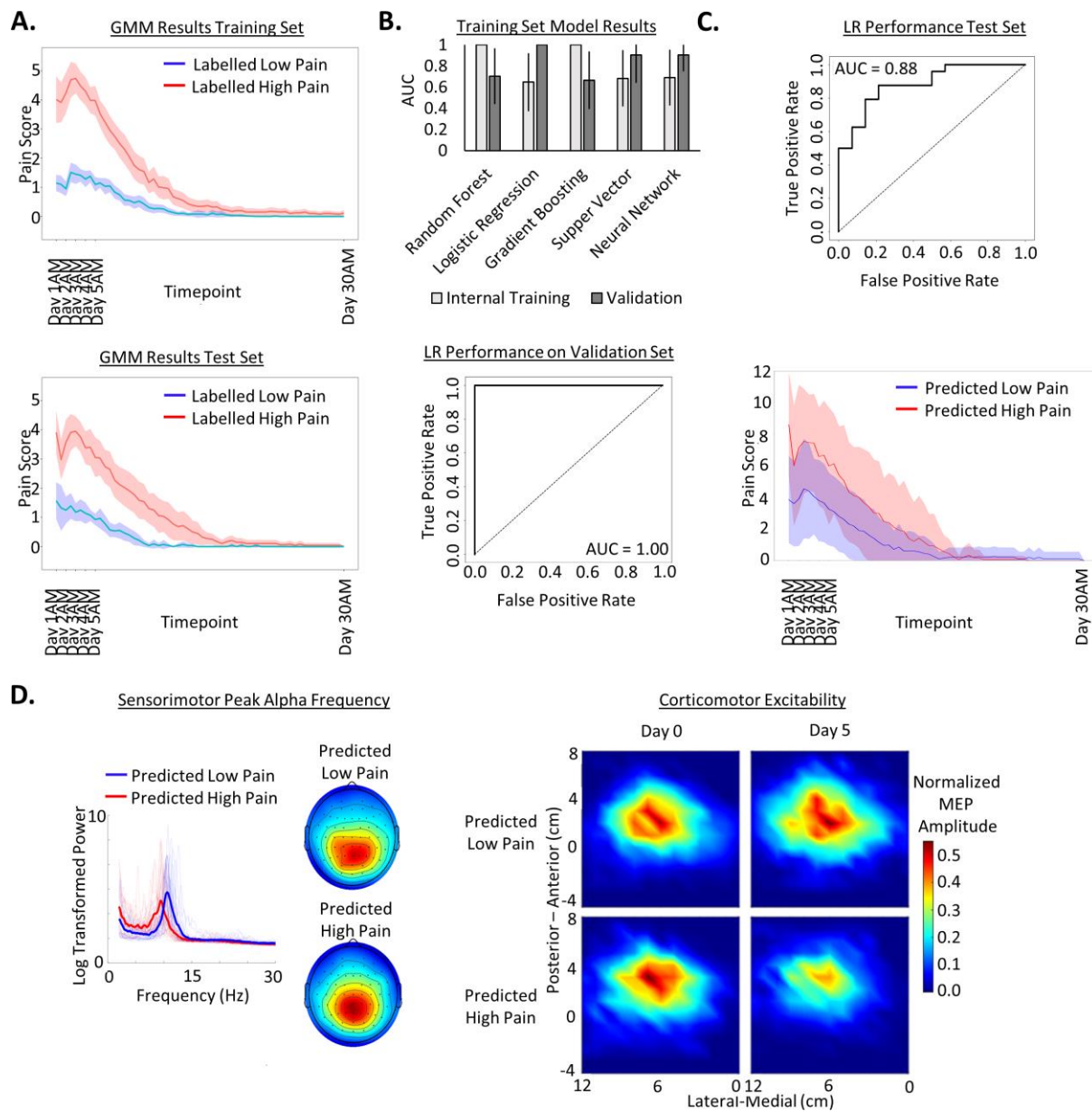
504

505

506

507

508 **Figure 2**



509

510

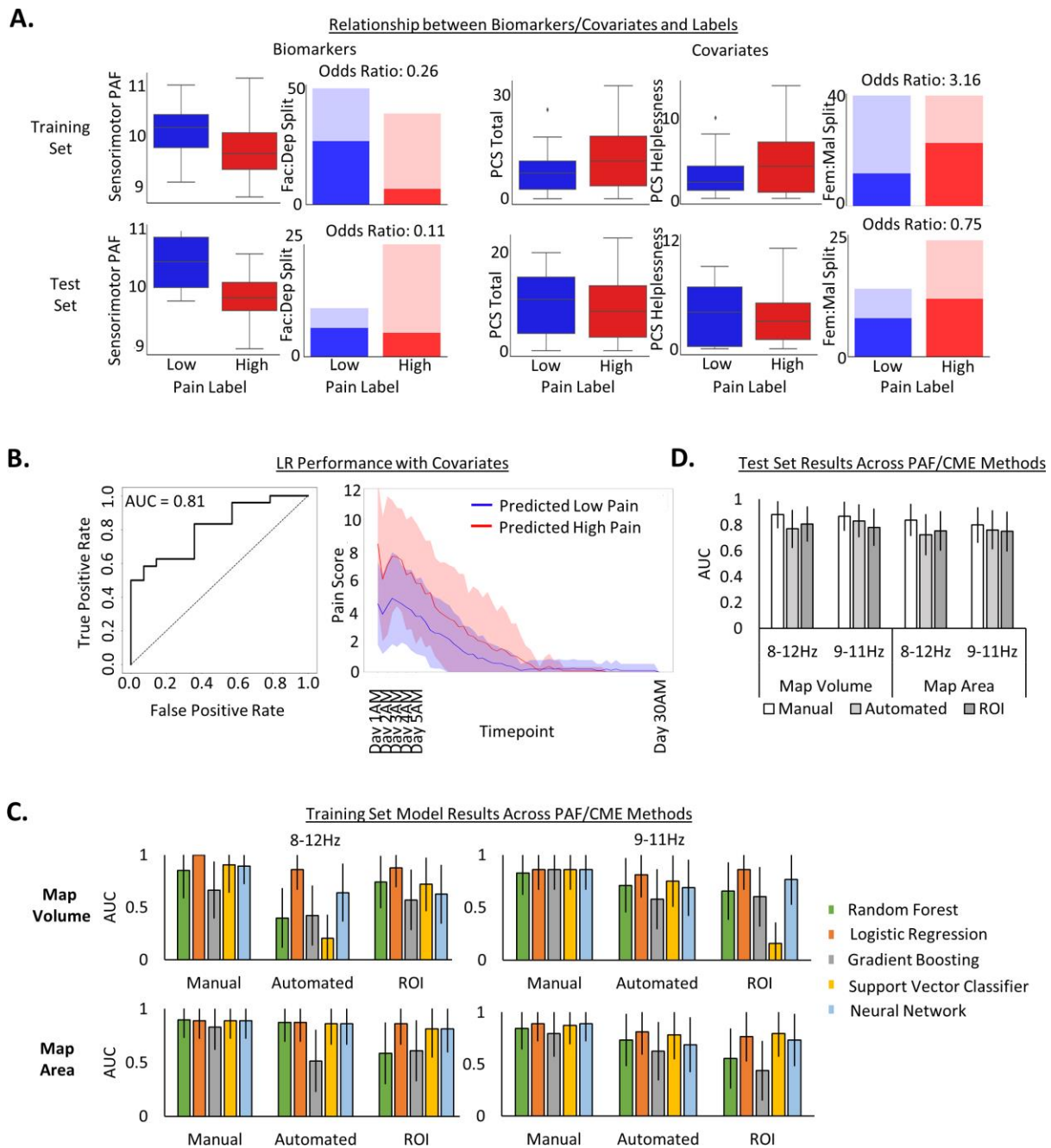
511

512

513

514

515 **Figure 3**



516

517

518

519

520