Prioritising Follow-Up for People with Suspected Epilepsy Using a Digital EEG Biomarker

Rosie Charles*[1], Emanuela De Falco*[1,2], Elizabeth Galizia[3], David Martin-Lopez[3], Kay Meiklejohn[4,1], David Allen[4], Lydia E. Staniaszek[5], Chris Price[6], Sophie Georgiou[6], Manny Bagary[7], Sakh Khalsa[7], Charlotte Lawthom[8], Rohit Shankar[9], John Terry[†1,2], Wessel Woldman[†1,2]

1. Neuronostics, Bristol, United Kingdom
2. University of Birmingham, Birmingham, United Kingdom
3. St George's Hospital NHS Foundation Trust, London, United Kingdom
4. University Hospital Southampton NHS Foundation Trust, Southampton, United Kingdom
5. University Hospitals Bristol and Weston NHS Foundation Trust, Bristol, United Kingdom
6. Royal Devon and Exeter NHS Foundation Trust, Exeter, United Kingdom
7. Birmingham and Solihull Mental Health NHS Foundation Trust, Birmingham, United Kingdom
8. Royal Gwent Hospital, Newport, United Kingdom
9. University of Plymouth, Plymouth, United Kingdom

**These authors contributed equally to this work as first authors*
†These authors contributed equally to this work as last authors

Abstract

Lengthy waits for follow-up testing are common for people with suspected epilepsy. This delays diagnosis, prolongs uncertainty and increases seizure risk. Initial EEGs are frequently inconclusive, yet follow-ups are typically dictated by referral date, rather than risk. Here, we tested whether a digital EEG biomarker could help prioritise those most likely to have epilepsy for expedited EEG testing. We analysed 196 non-contributory (non-diagnostic) initial EEGs collected from six National Health Service (NHS) sites in England. From these recordings, we extracted eight computational features and derived a digital biomarker that quantifies the likelihood of the EEG was recorded from someone with active epilepsy. We used this information to reorder follow-up lists and compared outcomes against standard referral-based scheduling. We found that ordering for follow-up testing based upon the digital biomarker consistently prioritised people subsequently diagnosed with epilepsy. The diagnostic yield of each subsequent EEG performed was increased relative to orderings based on time of referral. Our study indicates that a routine EEG may furnish an objective risk metric that could accelerate second-line investigations and so reduce diagnostic delay whilst improving resource allocation in clinical practice.

Key Words: **Digital biomarker; EEG analysis; Epilepsy Diagnosis; Machine Learning; Clinical Decision Support; Diagnostic Yield**

Introduction

An audit of UK and Ireland seizure referral pathways[1] found that though routine EEGs are performed in over 50% of first seizure patients, the diagnostic yield of these EEGs is limited. Another report highlighted a median sensitivity of just 17.3%[2]. The unfortunate reality: when a suspected first seizure triggers a specialist referral, an inconclusive EEG typically follows. From there, whilst individual hospital protocols vary, people are largely allocated for further testing based on time of referral. Of those receiving a second EEG, 17-34% are eventually diagnosed with epilepsy following ambulatory EEG[3], compared to 29% for sleep-deprived EEG[4].

However, limited resources often result in long waits for further testing creating significant diagnostic delays and substantial risks. A quarter of adults who experience a first unprovoked seizure will experience another within six months, rising to 35% by twelve months[5]. A retrospective study on newly diagnosed focal epilepsy patients at a Finnish epilepsy centre showed that the median diagnostic delay was 12 months, but could extend to several years, with 77% experiencing more than two further seizures before treatment began[6]. The consequences of diagnostic delay extend beyond immediate seizure risk, affecting driving privileges, employment security, psychological wellbeing, and potentially compromising long-term prognosis through delayed access to appropriate treatment.

Recently, computational approaches have shown promise in enhancing diagnostic yield. An Australia-based study found that scheduling video-EEG based on seizure-diary analysis significantly improved diagnostic yield from 50% to 62% ($p < 0.001$, >5000 video-EEGs), with confirmed seizures observed in 18% of high-risk cases versus 11% in standard scheduling[7].

Whilst seizure diaries offer valuable insights, their subjective nature is well documented[8]. More objective approaches have emerged through computational analysis of EEG data itself[9–14]. We have uncovered a set of eight spectral-, network- and computer model-based features, otherwise hidden to standard visual EEG analysis, which combine into a single digital biomarker of epilepsy: BioEP. In previous work, BioEP achieved a sensitivity of over 60% for detecting epilepsy from non-contributory EEGs[9]. BioEP's 'similarity score' compares the eight features calculated from an individual's non-contributory routine EEG with a database of people with known final diagnoses. Higher similarity scores indicate greater likelihood that the patient has epilepsy.

In this study, we investigated a practical application for this biomarker: prioritisation of patients for follow-up EEG testing based on seizure risk as guided by BioEP's 'similarity score'.

Using 196 non-contributory first routine EEGs collected across six first seizure clinics, we evaluate whether score-guided prioritisation improves diagnostic yield compared to

time-of-referral allocation. Our findings suggest that this approach ensures people ultimately diagnosed with epilepsy are seen sooner. This-pragmatic integration of digital innovation with clinical practice potentially enhances diagnostic efficiency, reduces risks of uncontrolled seizures, and lowers costs associated with delayed diagnosis.

Methods

Dataset

196 normal non-contributory first routine EEGs from six NHS sites were selected from a dataset previously described in Tait et al.[9]. Abnormal non-contributory EEGs, such as those including subtle abnormalities such as focal slowing, asymmetries and background disorganization were not included. In total 66 subjects were ultimately diagnosed with epilepsy ('Ep', both focal and generalised) and 130 subjects with an alternate condition ('NEp', including syncope or functional seizures). This distribution is consistent with the typical prevalence of epilepsy in individuals undergoing a second EEG[3].

Analysis

All analyses were performed using custom Python scripts. Pre-processing of EEG was performed using MNE-Python (v1.6.1)[15]. Statistical group comparisons were performed using the Wilcoxon signed-rank test from the scipy.stats module[16].

EEG preprocessing and biomarker extraction

EEGs were recorded using the international 10–20 system. We included 19 channels (Fp1, Fp2, F7, F3, Fz, F4, F8, C3, Cz, C4, P3, Pz, P4, T3, T4, T5, T6, O1, O2). EEGs were segmented into 20-second epochs, and epochs containing artefacts were excluded. Eight features were extracted, including spectral properties, network connectivity measures, and two model-based properties derived from mathematical frameworks described in previous work[9].

Confounder Model

Following previous findings[9], we accounted for confounding effects of sex and anti-seizure medication (ASM) status. For each extracted feature, confounder correction was applied using β coefficients derived from generalised linear models (GLMs) fitted to the training sets:

$$Feature_i^{Adj} = Feature_i - \beta_1 \cdot Sex - \beta_2 \cdot ASMstatus$$

Here $Feature_i$ represents one of the eight features, *Sex* is a categorical variable (0: female; 1: male; 2: other) and *ASMstatus* is a binary variable (0: not taking ASMs; 1: taking ASMs at time of EEG).

Prioritisation Based on Classifier Confidence

A random undersampling boosting (RUB) model with four principal components has previously been identified as the best-performing classifier of epilepsy[9]. For this study, we used a fixed partition: 106 EEGs (Ep = 36, NEp = 70) for training and 90 EEGs (Ep = 30, NEp = 60) for testing, with matching Ep:NEp ratios. This approach ensures that the EEG of each individual receives a single similarity score.

The trained RUB model produced a score between 0 and 1 for each individual in the testing set. This score represents the confidence of the classifier that a given EEG originated from an individual with active epilepsy, and as such is termed the "similarity score".

To simulate clinical prioritisation, we drew 50 random subsets of 40 individuals, each with a 50:50 Ep:NEp ratio. This balanced ratio was chosen as baseline scenario to evaluate our method. In each subset individuals were prioritised based on their classifier scores, from highest to lowest, to obtain an ordered list for follow-up. For comparison, we also generated a randomised order for each subset akin to standard referral-based scheduling.

We repeated this process for subsamples in which the prevalence of epilepsy varied from 10-80% to assess generalisability across different clinical settings.

Impact on Clinical metrics

We evaluated performance by comparing biomarker-ordered and random-ordered lists on several metrics. Specifically, we calculated the number of follow-up EEGs required to be completed to have seen 15%, 50%, and 80% of people with epilepsy. We also assessed cumulative percentage of Ep and NEp individuals, and the cumulative diagnostic yield for epilepsy as a function of follow-up EEGs performed on each ordered list.

To estimate diagnostic yield for epilepsy, we assumed an average sensitivity of 54% for second-line EEG investigations. This was estimated based on the sensitivity of the most common second-EEG types performed following a first routing EEG: ambulatory EEGs (sensitivity: 63%) and sleep-deprived EEGs (sensitivity: 45%)[17,18].

Yield was then calculated as:

$$Diagnostic\ yield = \frac{Number\ of\ Epilepsy\ patients\ seen\ \times\ Sensitivity}{Number\ of\ tests}$$

Results

We first assessed how effectively the biomarker prioritised people with epilepsy by estimating the number of EEGs required to have seen 15%, 50% and 80% of Ep patients across the 50 ordered subsamples (**Figure 1**). These values were compared to those obtained from the randomly ordered lists using the Wilcoxon signed-rank test. Biomarker-based ordering significantly reduced the number of EEGs needed to reach each threshold, indicating improved prioritisation of individuals with epilepsy.

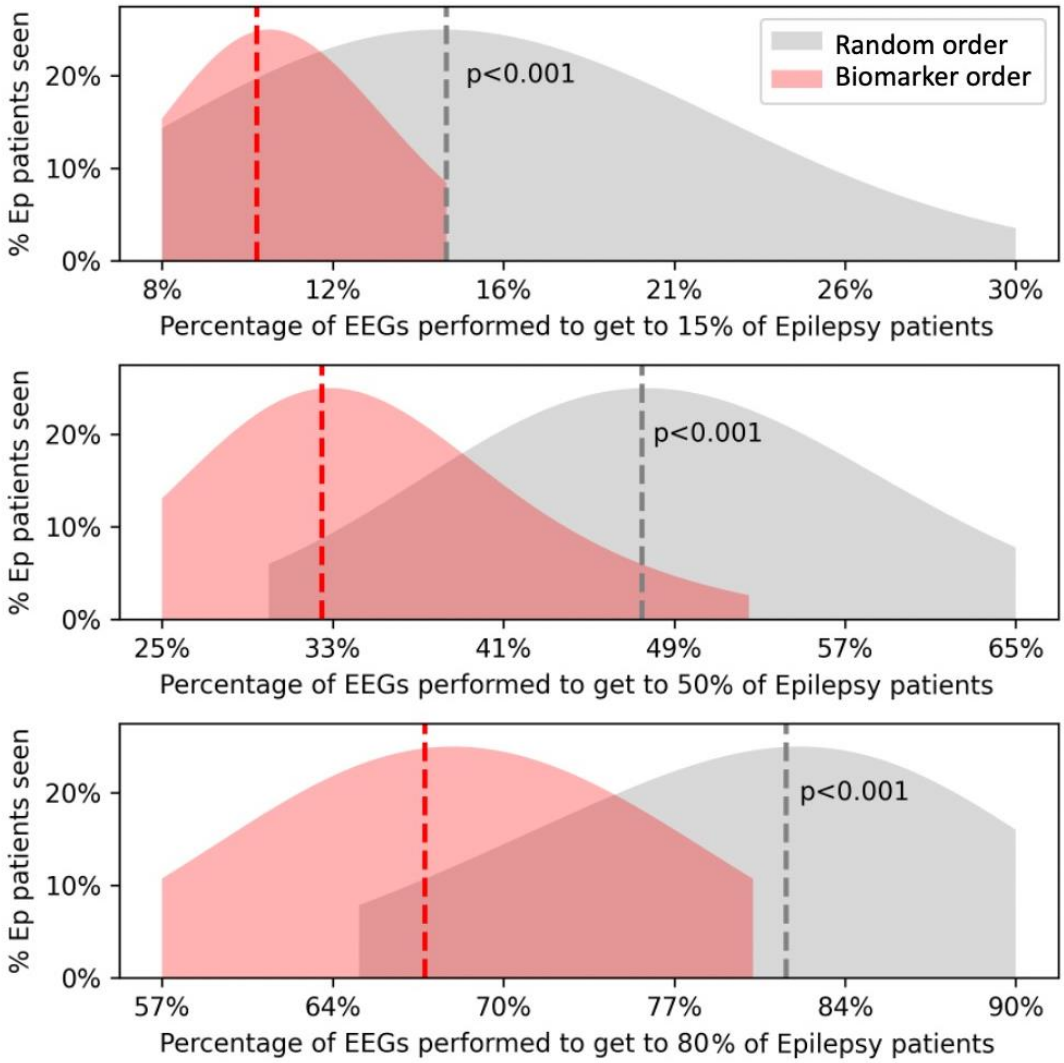

***Figure 1**. Numbers of EEGs needed to reach specific percentages of people with epilepsy receiving a follow-up EEG. Red curves show the distributions of EEGs required to reach 15%, 50% and 80% of Ep individuals across the 50 realisations. Grey curves represent the values obtained from the random distribution. Dotted vertical lines represent median of each distribution. p-values reflect statistical comparisons.*

To evaluate practical benefit of our re-ordering, we plotted the cumulative percentages of individuals seen (Ep and NEp) as a function of the percentage of EEGs performed (**Figure 2**.A1). This shows that our method successfully prioritises patients with epilepsy, with a consistently higher proportion of people with epilepsy appearing in the upper portion of a re-ordered list. In addition, we estimated the diagnostic yield for epilepsy following the biomarker-based reordering against random ordering (**Figure 2**.A2). Biomarker-based prioritisation consistently yielded improvements in diagnostic efficiency compared to random scheduling, above what would be expected by chance.

We next examined the midpoint impact: the number of people with epilepsy seen within the first 50% of the follow up visits (**Figure 2**.B1), and the corresponding diagnostic yield for epilepsy (**Figure 2**.B2). In both cases, biomarker-based ordering significantly outperformed random scheduling (Wilcoxon signed-rank test), highlighting the potential for earlier diagnosis and treatment.

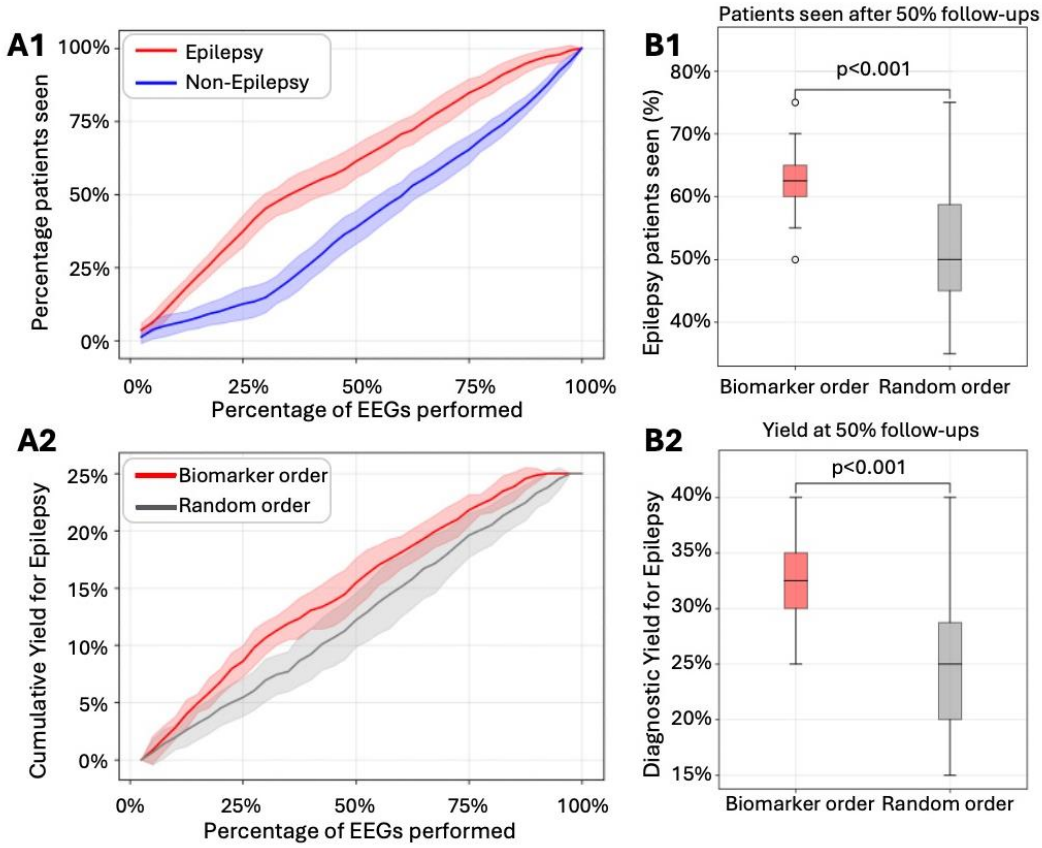

Figure 2. A1) *Percentage of Ep (red) and NEp (blue) individuals seen versus percentage of EEGs performed.* ***A2)*** *Diagnostic yield for epilepsy over time versus percentage of EEGs using the classifier ordering (red) and random scheduling (grey). Solid lines indicate mean value across the 50 realizations; shaded areas indicate standard deviations.* ***B1)*** *Number of people with epilepsy seen after 50% of the EEGs (biomarker-based (red) random scheduling (grey).* ***B2)*** *Corresponding diagnostic yield for epilepsy at 50% follow-up EEGs. Boxplots indicate median, and interquartile range. p-values reflect statistical comparisons.*

To explore robustness across different real-world scenarios, we repeated the above subsampling and reordering process using samples with varying epilepsy prevalence (10% to 80%) (**Figure S1**). The benefits of biomarker-based ordering were most pronounced when epilepsy prevalence was low to moderate (10–50%), which is closer to the typical prevalence observed in clinics. As expected, the marginal gain diminished as prevalence increased, given that a higher proportion of epilepsy cases results in a higher likelihood of capture by chance under random ordering.

Discussion

Significance

Our findings demonstrate that a digital EEG biomarker applied to non-contributory routine EEGs can be used to improve the prioritisation of follow-up testing for people with suspected epilepsy. By reordering individuals using a biomarker-derived similarity, more individuals with epilepsy are seen earlier, resulting in a higher diagnostic yield for epilepsy than seeing people based on time of referral. This has the potential to expedite diagnosis, reduce seizure burden, and improve resource allocation within stretched healthcare systems.

It must be noted that our study has inherent limitations that should be explored in a carefully designed prospective trial. Whilst the classifier was trained on hundreds of EEGs, expanding the training data to reflect broader, more heterogeneous, populations is necessary (in particular with respect to the rarer types of epilepsy). We also assumed uniform diagnostic yields for follow-up EEG types; real-world data will be important to confirm the robustness of this assumption. Within our cohort, diagnostic uncertainty was present given the non-contributory nature of the EEG. However, we cannot be certain that every individual (both Ep and NEp) had a follow-up EEG. Furthermore, our approach did not factor in existing clinical risks; whilst hybrid models could be developed in future studies that integrate both clinical risk factors as well as EEG based biomarkers, clinical decision making should always take precedence. Finally, the model was calibrated such that classification of those most likely to have epilepsy was prioritised. We should emphasise that individuals with non-epileptic causes of suspected seizures are of equal importance clinically.

Conclusion

In conclusion, integration of digital technologies into the diagnostic pathway offers a pragmatic, data-guided approach to addressing the challenge of epilepsy diagnosis in resource limited settings. As digital biomarkers continue to mature, they may support a more equitable and efficient use of EEG. Prioritisation of those most in need, without excluding others from follow-up should enhance and not replace clinical judgment.

Acknowledgements
W.W. was supported by Epilepsy Research UK (F2002), Innovate UK (103939), and the NIHR (AI01646). J.R.T. was supported by the EPSRC (EP/N014391/2 and EP/T027703/1), Innovate UK (103939), and the NIHR (AI01646). R.S. was supported by Innovate UK (103939) and the NIHR (AI01646). We thank Anke Verhaege and Amy Amin for comments on an earlier draft of the manuscript.

Funding Information
Funding was provided by Innovate UK, National Institute for Health and Care Research, Engineering and Physical Sciences Research Council, and Epilepsy Research UK.

Conflict of Interest Statement
R.C., E.D.F. and K.M. are employees of Neuronostics. W.W. and J.R.T. are cofounders, directors, and shareholders of Neuronostics. R.S. has received institutional research, travel support and/or honorarium for talks and expert advisory boards from LivaNova, UCB, Eisai, Veriton Pharma, Bial, Angelini, UnEEG and Jazz/GW pharma outside the submitted work. He holds or has held competitive grants from various national grant bodies including Innovate, Economic and Social Research Council (ESRC), Engineering and Physical Sciences Research Council (ESPRC), National Institute of Health Research (NIHR), NHS Small Business Research Initiative (SBRI) and other funding bodies including charities all outside this work. He has worked in a research capacity with Neuronostics as chief investigator for 3 projects including the project from which the data for this paper was derived.

None of the other authors have any conflict of interest to disclose. We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

Data Availability Statement
The EEG recordings and metadata are not publicly available due to restrictions by privacy laws. Postprocessed data supporting the findings of this study are available upon reasonable request from the corresponding author, and additional metadata may be made available upon reasonable request from the study sponsor (R.S.).

Ethics Statement
This study was approved by the HRA & HCRW (Integrated Research Application System (IRAS)24: 260729) and by the internal ethics committees of all collaborating clinical institutes.

References

1. Lee SH, Gillespie C, Bandyopadhyay S, Nazari A, Ooi SZY, Park JJ, et al. National audit of pathways in epileptic seizure referrals (NAPIER): A national, multicentre audit of first seizure clinics throughout the UK and Ireland. Seizure Eur J Epilepsy. 2023; 111:165–71.

2. Bouma HK, Labos C, Gore GC, Wolfson C, Keezer MR. The diagnostic accuracy of routine electroencephalography after a first unprovoked seizure. Eur J Neurol. 2015; 23(3):455–63.

3. Ray M, Sinopoulos N, Chandra L, Pang C. BSCN/ANS Ambulatory EEG Audit [Internet]. BSCN; 2019 [cited 2025]. Available from: https://www.bscn.org.uk/data/files/Audit/Audit18_ambulatory.pdf

4. Renzel R, Baumann CR, Poryazova R. EEG after sleep deprivation is a sensitive tool in the first diagnosis of idiopathic generalized but not focal epilepsy. Clin Neurophysiol. 2016; 127(1):209–13.

5. Neligan A, Adan G, Nevitt SJ, Pullen A, Sander JW, Bonnett L, et al. Prognosis of adults and children following a first unprovoked seizure. Cochrane Database Syst Rev. 2023; 2023(1):CD013847.

6. Parviainen L, Kälviäinen R, Jutila L. Impact of diagnostic delay on seizure outcome in newly diagnosed focal epilepsy. Epilepsia Open. 2020; 5(4):605–10.

7. Naim-Feil J, Stirling RE, Karoly PJ, Payne D, Winterling N, Eden D, et al. Pro-Ictal EEG Scheduling Improves the Yield of Epilepsy Monitoring: Validating the Use of Multiday Seizure Cycles to Optimize Video-EEG Timing. Ann Neurol. 2024; 96(6):1148–59.

8. Fisher RS, Blum DE, DiVentura B, Vannest J, Hixson JD, Moss R, et al. Seizure diaries for clinical research and practice: Limitations and future prospects. Epilepsy Behav. 2012; 24(3):304–10.

9. Tait L, Staniaszek LE, Galizia E, Martin-Lopez D, Walker MC, Azeez AAA, et al. Estimating the likelihood of epilepsy from clinically noncontributory electroencephalograms using computational analysis: A retrospective, multisite case–control study. Epilepsia. 2024; 65(8):2459–69.

10. Schmidt H, Woldman W, Goodfellow M, Chowdhury FA, Koutroumanidis M, Jewell S, et al. A computational biomarker of idiopathic generalized epilepsy from resting state EEG. Epilepsia [Internet]. 2016 [cited 2024]; 57(10). Available from: https://onlinelibrary.wiley.com/doi/10.1111/epi.13481

11. Woldman W, Schmidt H, Abela E, Chowdhury FA, Pawley AD, Jewell S, et al. Dynamic network properties of the interictal brain determine whether seizures appear focal or generalised. Sci Rep. 2020; 10(1):7043.

12. Myers P, Gunnarsdottir KM, Li A, Razskazovskiy V, Craley J, Chandler A, et al. Diagnosing Epilepsy with Normal Interictal EEG Using Dynamic Network Models. Ann Neurol. 2025; 97(5):907–18.

13. Lopez ET, Corsi M-C, Danieli A, Antoniazzi L, Angiolelli M, Bonanni P, et al. Dynamic reconfiguration of aperiodic brain activity supports cognitive functioning in epilepsy: A neural fingerprint identification. iScience [Internet]. 2025 [cited 2025]; 28(1). Available from: https://www.cell.com/iscience/abstract/S2589-0042(24)02724-X

14. Lopes MA, Krzemiński D, Hamandi K, Singh KD, Masuda N, Terry JR, et al. A computational biomarker of juvenile myoclonic epilepsy from resting-state MEG. Clin Neurophysiol. 2021; 132(4):922–7.

15. Larson E, Gramfort A, Engemann DA, Leppakangas J, Brodbeck C, Jas M, et al. MNE-Python [Internet]. Zenodo; 2024 [cited 2025]. Available from: https://zenodo.org/doi/10.5281/zenodo.592483

16. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. 2020; 17(3):261–72.

17. National Institute for Health and Care Excellence (NICE). Epilepsies in children, young people and adults - NICE Guideline [Internet]. NICE; 2022 [cited 2025]. Available from: https://www.nice.org.uk/guidance/ng217

18. Geut I, Weenink S, Knottnerus ILH, van Putten MJ a. M. Detecting interictal discharges in first seizure patients: ambulatory EEG or EEG after sleep deprivation? Seizure. 2017; 51:52–4.