

Computer-assisted analysis of routine EEG to identify hidden biomarkers of epilepsy: protocol for a systematic review

Authors

Émile Lemoine^{1, 2, 3}

Joel Neves Briard^{1, 3}

Bastien Rioux^{1, 3}

Renata Podbielski³

Bénédicte Nauche³

Denahin Toffa^{1, 3}

Mark Keezer^{1, 3, 5}

Frédéric Lesage²

Dang K. Nguyen^{1, 3}

Elie Bou Assi^{1, 3}

¹Department of Neurosciences, University of Montreal, Canada

²Institute of biomedical engineering, Polytechnique Montreal, Canada

³University of Montreal Hospital Center's Research Center, Canada

⁴School of Public Health, University of Montreal, Canada

⁵Stichting Epilepsie Instellingen Nederland (SEIN), Heemstede, The Netherlands

Corresponding author

Émile Lemoine, emile.lemoine@umontreal.ca

1051 rue Sanguinet, Montréal, Québec. H2X 3E4

Abstract

Background: The diagnosis of epilepsy frequently relies on the visual interpretation of the electroencephalogram (EEG) by a neurologist. The hallmark of epilepsy on EEG is the interictal epileptiform discharge (IED). This marker lacks sensitivity: it is only captured in a small percentage of 30-minute routine EEGs in patients with epilepsy. In the past three decades, there has been growing interest in the use of computational methods to analyze the EEG without relying on the detection of IEDs, but none have made it to the clinical practice. We aim to review the diagnostic accuracy of quantitative methods applied to ambulatory EEG analysis to guide the diagnosis and management of epilepsy.

Methods: The protocol complies with the recommendations for systematic reviews of diagnostic test accuracy by Cochrane. We will search MEDLINE, EMBASE, EBM reviews, IEEE Explore along with grey literature for articles, conference papers and conference abstracts published after 1961. We will include observational studies that present a computational method to analyze the EEG for the diagnosis of epilepsy in adults or children without relying on the identification of IEDs or seizures. The reference standard is the diagnosis of epilepsy by a physician. We will report the estimated pooled sensitivity and specificity, and receiver operating characteristic area-under-the-curve (ROC AUC) for each marker. If possible, we will perform a meta-analysis of the sensitivity and specificity and ROC AUC for each individual marker. We will assess the risk of bias using an adapted QUADAS-2 tool. We will also describe the algorithms used for signal processing, feature extraction and predictive modeling, and comment on the reproducibility of the different studies.

Discussion: Despite the promise to unveil epileptiform patterns that cannot be seen by the naked eye, computational analysis of ambulatory EEG has not yet been successfully translated to the clinical setting. We hope to produce recommendations for future studies on computer-assisted EEG interpretation for the diagnosis and management of epilepsy.

Systematic review registration: PROSPERO #292261

Keywords: Epilepsy – Electroencephalogram – Machine Learning – Diagnosis – Computer-assisted – Biomarker

1 **Background**

2 Epilepsy is characterized by an enduring propensity towards epileptic seizures—transient neurological
3 manifestations provoked by a state of abnormal and excessive neuronal activity in the brain¹. Epilepsy
4 affects over 65 millions of people worldwide, and 10% of the population will experience at least one
5 seizure in their lifetime^{2,3}. Epileptic seizures can lead to fractures, road accidents, isolation, anxiety,
6 cognitive decline, and death⁴. In specialized-care settings, the first anti-seizure medication (ASM)
7 achieves seizure freedom in approximately 47% of patients⁵. A prompt diagnosis is key in the prevention
8 of epilepsy-related morbidity and mortality⁴.

9 A history of epileptic seizures or a high recurrence risk after a single seizure are the basis for the
10 definition of epilepsy by the International League Against Epilepsy (ILAE)¹. Ancillary tests are often
11 needed to estimate seizure recurrence risk after a single seizure. These include the neurological
12 examination, neuroimaging, and the electroencephalogram (EEG).

13 An EEG records the electrical activity of the brain. It is recommended that all patients who present with a
14 first unprovoked seizure or with new diagnosis of epilepsy undergo an EEG^{6,7}. The initial EEG is
15 generally performed with electrodes applied to the patient's scalp (scalp EEG or *routine EEG*) for a
16 duration of 20–40 minutes⁸. The EEG tracing is then interpreted visually by a neurologist, who attempts
17 to identify interictal epileptiform discharges (IEDs; *aka* spikes). IEDs are brief (20–200ms) sharp
18 discharges, clearly emerging from background oscillations, often negative in polarity and sometimes
19 followed by a typical slow wave⁸. The presence of interictal spikes on the EEG is considered a hallmark
20 of epilepsy, as it represents a strong predictor of seizure recurrence^{9,10}. Furthermore, the identification of
21 interictal spikes can help localize an epileptic focus that may be amenable to surgical resection, and can
22 guide the withdrawal of ASMs in patients after a prolonged period of seizure freedom^{11,12}.

23 The interictal spike has several limitations. It occurs very sporadically: in patients with epilepsy, only 29
24 – 55% of routine EEGs will capture these transient abnormalities⁸. After a first unprovoked seizure in

25 adults, the sensitivity of a single routine EEG for detecting epileptiform abnormalities is only 17%⁹.
26 Furthermore, their identification is somewhat subjective: the percent agreement between EEG experts is
27 around 76%¹³. Many physiological transient discharges¹³ can be misinterpreted as epileptiform spikes. This
28 can lead to the erroneous diagnosis of epilepsy, with sometimes important consequences^{14,15}. In patients
29 labelled with drug-resistant epilepsy, over 25% may have had an erroneous diagnosis as a result of both
30 inadequate history taking and misinterpretation of the EEG¹⁶. Despite the abundant information on brain
31 activity recorded by the EEG, no other interictal anomalies have been validated for use in clinical
32 settings^{1,17,18}.

33 Compared to other neuroimaging modalities, a scalp EEG is inexpensive, easy to acquire, and confers
34 functional information with high temporal resolution^{19,20}. Moreover, great effort was put in the last decade
35 by the ILAE in standardizing the equipment, recording and storage of EEG data^{10,21}. Decades of research
36 have demonstrated that the automated analysis of EEG can identify hidden differences between with
37 epilepsy and non-epileptic subjects in terms of connectivity²²⁻²⁴, signal predictability and complexity^{25,26},
38 spectral power^{27,28}, and chaoticity²⁹. Computational analysis of EEG holds the promise of extracting
39 information that is invisible to the naked eye of the human interpreter, in an objective and reproducible
40 manner. Discovering new, non-visible markers of epilepsy could increase the diagnostic yield of the EEG,
41 improve its accessibility, and reduce costs, especially in settings where the expertise of a fellowship-
42 trained neurophysiologist is unavailable^{18,30}. In spite of this, none of the proposed non-visible markers of
43 epilepsy have made it into clinical practice^{10,31}.

44 We will perform a systematic review of diagnostic test accuracy for automated methods of EEG analysis
45 to distinguish between patients with and without epilepsy without relying on the detection of spikes and
46 seizures. The questions that this review addresses are the following: What is the current evidence on the
47 performances of automatically extracted hidden markers of epilepsy for the diagnosis of epilepsy? And
48 what are the different algorithms that have been tested and how does their diagnostic accuracy compare?

49 **Methods**

50 **Study design**

51 This will be a systematic review and meta-analysis following guidance from the Cochrane Diagnostic
52 Test Accuracy group. We will report the results according to the PRISMA statement for diagnostic test
53 accuracy (PRISMA-DTA)³².

54 **Study selection criteria**

55 **Type of studies**

56 We will include all studies that describe a computed marker of epilepsy on routine (scalp) EEG which
57 does not explicitly rely on the identification of interictal spikes or ictal activity (seizures). Studies must
58 compare the EEG signal of individuals with and without epilepsy. We will include retrospective or
59 prospective comparative studies enabling the assessment of diagnostic accuracy (cohort or case-control
60 studies). We will exclude studies reporting data on non-human animals only, studies that include only
61 intracranial or critical care EEG recordings, studies that do not include both individuals with and without
62 epilepsy, and studies that are focused solely on seizure/spike detection or on short-term (<24h) seizure
63 prediction. For studies that include multiple EEG types, we will only extract data that meet the inclusion
64 criteria. We restricted the search to studies published after 1961 (the first use of digital EEG)³³. There are
65 no restrictions for language.

66 **Population**

67 Our population of interest is individuals undergoing routine EEG in a clinical or research setting. A
68 routine EEG is defined as a 20- to 60-minute scalp recording using the international 10–20 electrodes
69 system, with or without prior sleep deprivation. There is no restriction for age groups or diagnoses.

70 **Reference standard**

71 We defined the reference standard as the diagnosis of epilepsy by a physician based on criteria specified
72 by the authors (clinical or para-clinical). These criteria must accord with the definition of epilepsy by the
73 ILAE: having had at least one seizure and long-term enduring predisposition to other unprovoked
74 seizures^{1,34}.

75 **Index test**

76 The index test is a characteristic or feature which is computationally extracted from the EEG signal to
77 identify patients with epilepsy, without relying on detecting IEDs or seizures. These include measures of
78 connectivity, entropy, chaoticity, and power spectrum density³⁵. Also included are statistical models that
79 combine several features or models that take as input the raw or processed EEG.

80 **Search strategy**

81 The search strategy (**Appendix 1**) was developed by two medical librarians specialized in systematic
82 reviews (BN and RP), and peer-reviewed by a senior colleague. We will search MEDLINE (Ovid),
83 EMBASE (Ovid), EBM reviews (Ovid), IEEE Explore along with grey literature for articles, conference
84 papers and conference abstracts. We will use the Covidence platform (Melbourne, Australia) to manage
85 our data for eligibility assessment, selection, and data collection. Two independent reviewers (EL, and
86 either JNB or BR) will screen the records for eligibility using their title and abstract. Any item selected by
87 either reviewer will proceed to the next phase. This process will be repeated on the screened items, this
88 time by consulting the items' full text. A third, senior reviewer (EBA) will settle conflicts as necessary
89 during the final selection.

90 **Data items**

91 Data collection will be performed using Covidence by two independent reviewers (EL and JNB/BR), and
92 conflicts will be resolved by a third author (EBA). Authors of the primary study will be contacted if the

93 required data are not available in the original publication. Data collection will include the following
94 information:

- 95 1. Title and authors of the study, country of sampling, year of publication;
- 96 2. Study type: retrospective vs. prospective, design (cohort, case control);
- 97 3. Study sample: exclusion and inclusion criteria, number of screened and included patients;
- 98 4. Data collection:
 - 99 a. Number of patients, number of EEGs, duration of EEG recordings, use of activation
100 procedures (hyperventilation, photic stimulation, sleep deprivation), setting of recording
101 (hospitalized or ambulatory), whether the same protocol was used for all patients;
 - 102 b. Number of electrodes, sampling frequency;
 - 103 c. If public dataset: reference to the original dataset, dataset name, exclusion/inclusion
104 criteria used on the EEG segments from the dataset;
 - 105 d. Participant characteristics: age, sex, comorbidities, number of ASM, age of first seizure;
- 106 5. Reference standard: whether a predefined reference standard was used, definition of reference
107 standard, whether all patients underwent the same reference standard, time lapse between
108 reference standard and EEG;
- 109 6. Index test:
 - 110 a. Pre-processing: artifact detection and removal (automated or manual), filtering method,
111 filtering frequencies, segmentation protocol (whole EEG vs. EEG segments, window
112 size, overlapping vs. non-overlapping segments, manual vs. automated selection of
113 segments), channel selection;
 - 114 b. Feature extraction and selection: multi-channel vs. single channel, number of channels
115 selected, whether feature selection was performed, feature extraction algorithm, feature
116 selection method, whether feature selection was applied to data before vs. after excluding
117 testing data;

- 118 c. Classification: algorithm(s) used for classification, testing methodology (cross-validation
119 vs. held out testing set);
- 120 d. Metric used to report diagnostic performances: ROC AUC,
121 accuracy/sensitivity/specificity, F₁-score, reporting of confidence intervals (CI);
- 122 7. Diagnostic performances: number of true positives, number of true negatives, number of false
123 positives, number of false negatives, reported accuracy, reported sensitivity, reported specificity,
124 reported F₁-score, reported ROC AUC (if more than one index test is performed on the same
125 patient, we will only consider the first test);
- 126 8. Reproducibility: whether every data processing step is detailed, whether methods can be
127 reproduced easily, data availability, code availability, open-source computer libraries referenced.

128 **Risk of bias**

129 The risk of bias of all included studies will be assessed through an adapted version of the QUADAS-2
130 tool³⁶. Risk of bias for each of the following four elements will be evaluated by two independent
131 reviewers (EL and JNB/BR) as low, high, or unclear. Conflicts will be resolved by a third author (EBA).
132 In addition, all publicly available datasets used by at least one of the included studies will be evaluated
133 with the same tool. The following items will be assessed:

- 134 1. Patient selection
- 135 a. Is the population representative of clinical practice?
- 136 b. Are inclusion and exclusion criteria identical for cases (patients with epilepsy) and
137 controls?
- 138 c. Are withdrawals explained and appropriate? If individual EEG segments were excluded,
139 were the same criteria used for all segments?
- 140 2. Index test
- 141 a. Were the protocols used for recording the EEG identical in all patients, irrespective of the
142 epilepsy diagnosis?

- 143 b. Was the index test validated on an independent sample of patients (patients which were
144 not used to identify the index test's threshold or train the learning algorithm)?
- 145 3. Reference standard
- 146 a. Are the criteria used for the diagnosis of epilepsy specified and acceptable (likely to
147 correctly classify the target condition)?
- 148 b. Was the reference standard assessment independent and blinded to the index test?
- 149 4. Flow and timing
- 150 a. Did the whole sample undergo the reference standard?
- 151 b. Did the whole sample undergo the same reference standard?
- 152 c. Was the time lapse between reference standard and EEG acceptable?
- 153 d. Was the same data used in the index method available at the time of the reference
154 standard?
- 155 e. Were all EEGs included in the analysis?

156 **Data synthesis**

157 We will provide a table summarizing every published study included in the review, comparing the
158 studies' design, population, reference standard, dataset size, data processing methods, and diagnostic
159 accuracy. We will also provide a table summarizing the risk of bias for all items in the adapted
160 QUADAS-2 tool, comparing 1) every individual article included in the review, and 2) every public
161 dataset that is used in ≥ 2 studies.

162 We will describe the number of patients, number of EEGs, duration of EEGs, and the EEG-duration-per-
163 patient ratio across all included studies. We will report the pooled proportion of patients with focal vs.
164 generalized epilepsy, adult vs. children, structural vs. non-structural epilepsy, and with specific epilepsy
165 syndromes. For every publicly available dataset identified during the review, we will report the number of
166 studies that used that dataset in their work.

167 We will summarize the methods used by the different articles during the pipeline's algorithm (pre-
168 processing, feature extraction, feature selection, and classification algorithm), along with the proportion
169 of studies that used each method.

170 **Analyses**

171 We will estimate the specificity and sensitivity for each study, using the Wilson score to compute 95%
172 CI. For studies with varying thresholds, we will estimate the ROC AUC and 95% CI.

173 If there are sufficient (≥ 5) studies that report the number of true/false positives and true/false negatives,
174 we will estimate the pooled sensitivity and specificity of each individual marker using a hierarchical,
175 bivariate generalized linear mixed model³⁷. This allows us to account for the correlation between
176 specificity and sensitivity in a single study. If ≥ 5 studies report these numbers for varying thresholds, we
177 will estimate the pooled ROC curve using the Rutter and Gatsonis HSROC model³⁸. All analyses will be
178 implemented with the R statistical language. A p -value < 0.05 will be considered statistically significant.
179 Given insufficient data for the pooled estimates, we will only describe the diagnostic performances
180 (sensitivity, specificity, ROC AUC) narratively. We will present the results of the analyses with forest
181 plots.

182 We will quantify heterogeneity using the variances of the logit specificity and sensitivity, as well as the
183 median odds ratio (median OR)³⁹. The median OR is a measure of inter-study variance translated on the
184 OR scale. It corresponds to the increase in the odds of being true positive/negative in a patient/control
185 going from a study with lower sensitivity/specificity to a study with higher sensitivity/specificity. For
186 heterogeneity in the ROC plane, we will compute the area of the 95% prediction ellipse³⁹. The median OR
187 and the area of the 95% prediction ellipse are easily obtained and interpreted, and take into account the
188 correlation between a single study's specificity and sensitivity in contrast to univariate methods like
189 Cochran's Q and I^2 ^{37,40}. We will perform subgroup analysis for the following variables: epilepsy type
190 (focal, generalized), epilepsy etiology (structural vs. non-structural), age groups (children (< 18 y.o.),
191 adults (≥ 18 y.o.)), epilepsy syndromes, extracted marker, and dataset used. We will assess heterogeneity

192 for all subgroup analyses. We will consider a study as belonging to a particular subgroup if $\geq 80\%$ of the
193 studied population belongs to that subgroup. Sensitivity analysis will be conducted for the main analyses
194 by excluding studies with overall high/unclear risk of bias.

195 Some studies use multiple markers to classify patients with epilepsy from controls (*e.g.*, as input
196 features for a machine learning algorithm). For each marker that is used in ≥ 2 of such studies, we will
197 evaluate the number of studies for which these markers were identified as “important” (selected for the
198 classification task or statistically significant in separating the two classes) and the ratio between the
199 number of studies in which this marker was extracted vs. identified as important.

200 Reporting bias for sensitivity and specificity will be evaluated by visual inspection of funnel plots.

201 **Discussion**

202 The interictal EEG is key in the diagnosis of epilepsy, solely based on the visual identification of
203 interictal spikes⁴¹. Despite years of research on computational biomarkers of epilepsy, only these spikes
204 are currently used in clinical settings^{1,17,18}. This review aims to systematically evaluate the diagnostic
205 performances of hidden interictal markers of epilepsy on EEG, describe the data processing pipelines
206 favored by the researchers to classify the EEG for epilepsy diagnosis, and identify the pitfalls that prevent
207 clinical translation of these algorithms.

208 Algorithms have gained growing interest in medicine for their potential to assist diagnosis and guide
209 clinical decision-making⁴². EEG analysis is well-suited for this application due to the complex nature of
210 the EEG signal. Automated extraction of new epilepsy markers on routine EEG could lead to reduced rate
211 of misdiagnosis, increased availability in areas without access to an expert neurophysiologist, and more
212 efficient clinical trials. Research on automatic analysis of EEG data is thriving, in part assisted by the
213 recent increase in computational capacities⁴³⁻⁵⁰. However, automatic analysis of EEG is not mentioned in
214 any of the high-quality clinical practice guidelines systematically reviewed by the ILAE¹⁷.

215 In recent years, increased computational capacities have allowed the development of powerful algorithms
216 that can learn complex representations such as medical images and EEG signals^{43,51,52}. A growing number
217 of algorithms have now been approved by the United States Food and Drug Administration for assisting
218 in the diagnosis of several diseases⁵³. Recent systematic reviews have found that most of the studies on
219 automated diagnosis using artificial intelligence have high risk of bias, mostly due to patient selection
220 methodology and absence of validation on external data⁵⁴⁻⁵⁶. Systematic reviews on computer-based
221 clinical-decision support systems also highlight the need for more robust patient selection⁵⁷⁻⁶².

222 Translation of technology to clinical practice requires strong evidence based on high quality research.
223 This review is important because it will establish the potential of automatic analysis of EEG as a
224 diagnostic tool for epilepsy, and if evidence to support its use is lacking, it will identify the pitfalls that
225 need to be overcome in future research. Also, by systematically describing current practices that are used
226 by research groups, it will serve as a reference for new researchers in the field.

227 We anticipate that diagnostic accuracy of automatic analysis of EEG for epilepsy will be hard to estimate
228 because of the high heterogeneity between the different dataset used and between the data processing
229 methodology. We also anticipate high risk of bias in many studies, because of the high volume of “proof-
230 of-concept” studies that emphasize computation performances and algorithm development over rigorous
231 diagnostic study methodology. In these cases, we hope to produce recommendations that will assist in
232 bridging the gap between the development of new automated markers and validation in appropriate
233 populations, for ultimate implementation into clinical practice.

234 **List of abbreviations**

235 ASM: anti-seizure medication; CI: confidence interval; EEG: electroencephalogram; IED: interictal
236 epileptiform discharge; ILAE: International League Against Epilepsy; ROC AUC: receiver operating-
237 characteristic area-under-the-curve.

238 **Declarations**

239 MRK and DKN report unrestricted educational grants from UCB and Eisai, and research grants for
240 investigator-initiated studies from UCB and Eisai. Émile Lemoine is supported by a scholarship from the
241 Canadian Institute of Health Research. Dang Nguyen is supported by the Canada Research Chairs
242 Program, the Canadian Institutes of Health Research, and Natural Sciences and Engineering Research
243 Council of Canada. Data collected for this study will be available upon reasonable request.

244 **Authors' contributions**

245 EL planned the study, drafted the protocol, reviewed the search strategy, and is the guarantor of the
246 review. DT, FL, DKN, and EBA participated in the design of the study. JNB, BR, DT, MRK, FL, DKN,
247 and EBA provided content expertise and critically reviewed the manuscript and the search strategy. BN
248 and RP designed the search strategy. All authors read and approved the final manuscript.

249 **References**

- 250 1. Fisher, R. S. *et al.* ILAE Official Report: A practical clinical definition of epilepsy. *Epilepsia* **55**,
251 475–482 (2014).
- 252 2. Ngugi, A. K., Bottomley, C., Kleinschmidt, I., Sander, J. W. & Newton, C. R. Estimation of the
253 burden of active and life-time epilepsy: a meta-analytic approach. *Epilepsia* **51**, 883–890 (2010).
- 254 3. Hauser, W. A. & Beghi, E. First seizure definitions and worldwide incidence and mortality. *Epilepsia*
255 **49 Suppl 1**, 8–12 (2008).
- 256 4. Devinsky, O., Spruill, T., Thurman, D. & Friedman, D. Recognizing and preventing epilepsy-related
257 mortality: A call for action. *Neurology* **86**, 779–786 (2016).
- 258 5. Kwan, P. & Brodie, M. J. Early identification of refractory epilepsy. *N Engl J Med* **342**, 314–319
259 (2000).

- 260 6. Krumholz, A. *et al.* Practice Parameter: evaluating an apparent unprovoked first seizure in adults (an
261 evidence-based review): report of the Quality Standards Subcommittee of the American Academy of
262 Neurology and the American Epilepsy Society. *Neurology* **69**, 1996–2007 (2007).
- 263 7. Hirtz, D. *et al.* Practice parameter: evaluating a first nonfebrile seizure in children: report of the
264 quality standards subcommittee of the American Academy of Neurology, The Child Neurology
265 Society, and The American Epilepsy Society. *Neurology* **55**, 616–623 (2000).
- 266 8. Pillai, J. & Sperling, M. R. Interictal EEG and the Diagnosis of Epilepsy. *Epilepsia* **47**, 14–22 (2006).
- 267 9. Bouma, H. K., Labos, C., Gore, G. C., Wolfson, C. & Keezer, M. R. The diagnostic accuracy of
268 routine electroencephalography after a first unprovoked seizure. *European Journal of Neurology* **23**,
269 455–463 (2016).
- 270 10. Tatum, W. O. *et al.* Clinical utility of EEG in diagnosing and monitoring epilepsy in adults. *Clinical*
271 *Neurophysiology* **129**, 1056–1082 (2018).
- 272 11. Lamberink, H. J. *et al.* Individualised prediction model of seizure recurrence and long-term outcomes
273 after withdrawal of antiepileptic drugs in seizure-free patients: a systematic review and individual
274 participant data meta-analysis. *Lancet Neurol* **16**, 523–531 (2017).
- 275 12. West, S. *et al.* Surgery for epilepsy. *Cochrane Database of Systematic Reviews* (2019)
276 doi:10.1002/14651858.CD010541.pub3.
- 277 13. Jing, J. *et al.* Interrater Reliability of Experts in Identifying Interictal Epileptiform Discharges in
278 Electroencephalograms. *JAMA Neurology* **77**, 49–57 (2020).
- 279 14. Amin, U. & Benbadis, S. R. The Role of EEG in the Erroneous Diagnosis of Epilepsy. *J Clin*
280 *Neurophysiol* **36**, (2019).
- 281 15. Kang, J. Y. & Krauss, G. L. Normal Variants Are Commonly Overread as Interictal Epileptiform
282 Abnormalities. *J Clin Neurophysiol* **36**, 257–263 (2019).
- 283 16. Smith, D., Defalla, B. A. & Chadwick, D. W. The misdiagnosis of epilepsy and the management of
284 refractory epilepsy in a specialist clinic. *QJM* **92**, 15–23 (1999).

- 285 17. Sauro, K. M. *et al.* The current state of epilepsy guidelines: A systematic review. *Epilepsia* **57**, 13–23
286 (2016).
- 287 18. Engel Jr, J., Bragin, A. & Staba, R. Nonictal EEG biomarkers for diagnosis and treatment. *Epilepsia*
288 *Open* **3**, 120–126 (2018).
- 289 19. DellaBadia Jr, J., Bell, W. L., Keyes Jr, J. W., Mathews, V. P. & Glazier, S. S. Assessment and cost
290 comparison of sleep-deprived EEG, MRI and PET in the prediction of surgical treatment for epilepsy.
291 *Seizure* **11**, 303–309 (2002).
- 292 20. Abdelhady, S., Shokri, H., Fathy, M. & wahid el din, mona M. Evaluation of the direct costs of
293 epilepsy in a sample of Egyptian patients following up in Ain Shams University Hospital. *The*
294 *Egyptian Journal of Neurology, Psychiatry and Neurosurgery* **56**, 112 (2020).
- 295 21. Velis, D., Plouin, P., Gotman, J., Da Silva, F. L., & members of the ILAE DMC Subcommittee on
296 Neurophysiology. Recommendations Regarding the Requirements and Applications for Long-term
297 Recordings in Epilepsy. *Epilepsia* **48**, 379–384 (2007).
- 298 22. Schmidt, H. *et al.* A computational biomarker of idiopathic generalized epilepsy from resting state
299 EEG. *Epilepsia* **57**, e200–e204 (2016).
- 300 23. Lopes, M. A. *et al.* Revealing epilepsy type using a computational analysis of interictal EEG.
301 *Scientific Reports* **9**, 10169 (2019).
- 302 24. Verhoeven, T. *et al.* Automated diagnosis of temporal lobe epilepsy in the absence of interictal
303 spikes. *NeuroImage: Clinical* **17**, 10–15 (2018).
- 304 25. Ouyang, C.-S., Yang, R.-C., Wu, R.-C., Chiang, C.-T. & Lin, L.-C. Determination of Antiepileptic
305 Drugs Withdrawal Through EEG Hjorth Parameter Analysis. *Int. J. Neur. Syst.* **30**, 2050036 (2020).
- 306 26. Zhang, J.-H. *et al.* Personalized prediction model for seizure-free epilepsy with levetiracetam therapy:
307 a retrospective data analysis using support vector machine. *Br J Clin Pharmacol* **84**, 2615–2624
308 (2018).

- 309 27. Oliva, J. T. & Rosa, J. L. G. Differentiation between Normal and Interictal EEG Using Multitaper
310 Spectral Classifiers. in *2018 International Joint Conference on Neural Networks (IJCNN)* 1–8
311 (2018). doi:10.1109/IJCNN.2018.8489503.
- 312 28. Pegg, E. J., Taylor, J. R. & Mohanraj, R. Spectral power of interictal EEG in the diagnosis and
313 prognosis of idiopathic generalized epilepsies. *Epilepsy & Behavior* **112**, 107427 (2020).
- 314 29. Jacob, J. E., Sreelatha, V. V., Iype, T., Nair, G. K. & Yohannan, D. G. Diagnosis of epilepsy from
315 interictal EEGs based on chaotic and wavelet transformation. *Analog Integrated Circuits and Signal*
316 *Processing* **89**, 131–138 (2016).
- 317 30. Wahl, B., Cossy-Gantner, A., Germann, S. & Schwalbe, N. R. Artificial intelligence (AI) and global
318 health: how can AI contribute to health in resource-poor settings? *BMJ Global Health* **3**, e000798
319 (2018).
- 320 31. Pitkänen, A. *et al.* Advances in the development of biomarkers for epilepsy. *The Lancet Neurology*
321 **15**, 843–856 (2016).
- 322 32. McInnes, M. D. F. *et al.* Preferred Reporting Items for a Systematic Review and Meta-analysis of
323 Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. *JAMA* **319**, 388–396 (2018).
- 324 33. November, J. Biomedical computing: Digitizing life in the United States. *Biomedical Computing:*
325 *Digitizing Life in the United States* 1–344 (2012).
- 326 34. Fisher, R. S. *et al.* Epileptic seizures and epilepsy: definitions proposed by the International League
327 Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia* **46**, 470–472
328 (2005).
- 329 35. Supriya, S., Siuly, S., Wang, H. & Zhang, Y. Automated epilepsy detection techniques from
330 electroencephalogram signals: a review study. *Health Information Science and Systems* **8**, 33 (2020).
- 331 36. Whiting, P. F. *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy
332 studies. *Ann Intern Med* **155**, 529–536 (2011).
- 333 37. Reitsma, J. B. *et al.* Bivariate analysis of sensitivity and specificity produces informative summary
334 measures in diagnostic reviews. *Journal of Clinical Epidemiology* **58**, 982–990 (2005).

- 335 38. Rutter, C. M. & Gatsonis, C. A. A hierarchical regression approach to meta-analysis of diagnostic test
336 accuracy evaluations. *Stat Med* **20**, 2865–2884 (2001).
- 337 39. Plana, M. N., Pérez, T. & Zamora, J. New measures improved the reporting of heterogeneity in
338 diagnostic test accuracy reviews: a metaepidemiological study. *Journal of Clinical Epidemiology*
339 **131**, 101–112 (2021).
- 340 40. Rücker, G., Schwarzer, G., Carpenter, J. R. & Schumacher, M. Undue reliance on I² in assessing
341 heterogeneity may mislead. *BMC Medical Research Methodology* **8**, 79 (2008).
- 342 41. Smith, S. J. M. EEG in the diagnosis, classification, and management of patients with epilepsy. *J*
343 *Neurol Neurosurg Psychiatry* **76**, ii2–ii7 (2005).
- 344 42. Obermeyer, Z. & Emanuel, E. J. Predicting the Future - Big Data, Machine Learning, and Clinical
345 Medicine. *N Engl J Med* **375**, 1216–1219 (2016).
- 346 43. Roy, Y. *et al.* Deep learning-based electroencephalography analysis: a systematic review. *Journal of*
347 *Neural Engineering* **16**, 051001 (2019).
- 348 44. Craik, A., He, Y. & Contreras-Vidal, J. L. Deep learning for electroencephalogram (EEG)
349 classification tasks: a review. *J. Neural Eng.* **16**, 031001 (2019).
- 350 45. Rasheed, K. *et al.* Machine Learning for Predicting Epileptic Seizures Using EEG Signals: A
351 Review. *IEEE Rev. Biomed. Eng.* **14**, 139–155 (2020).
- 352 46. Gemein, L. A. W. *et al.* Machine-learning-based diagnostics of EEG pathology. *NeuroImage* **220**,
353 117021 (2020).
- 354 47. Mesraoua, B. *et al.* Electroencephalography in epilepsy: look for what could be beyond the visual
355 inspection. *Neurological Sciences* **40**, 2287–2291 (2019).
- 356 48. van Diessen, E. *et al.* Brain Network Organization in Focal Epilepsy: A Systematic Review and
357 Meta-Analysis. *PLOS ONE* **9**, e114606 (2014).
- 358 49. Faiman, I., Smith, S., Hodsoll, J., Young, A. H. & Shotbolt, P. Resting-state EEG for the diagnosis of
359 idiopathic epilepsy and psychogenic nonepileptic seizures: A systematic review. *Epilepsy & Behavior*
360 **121**, 108047 (2021).

- 361 50. Pegg, E. J., Taylor, J. R., Keller, S. S. & Mohanraj, R. Interictal structural and functional connectivity
362 in idiopathic generalized epilepsy: A systematic review of graph theoretical studies. *Epilepsy &*
363 *Behavior* **106**, (2020).
- 364 51. Esteva, A. *et al.* A guide to deep learning in healthcare. *Nature Medicine* **25**, 24–29 (2019).
- 365 52. Litjens, G. *et al.* A survey on deep learning in medical image analysis. *Medical Image Analysis* **42**,
366 60–88 (2017).
- 367 53. FDA Cleared AI Algorithms. *American College of Radiology Data Science Institute*
368 <https://models.acrdsi.org>.
- 369 54. Aggarwal, R. *et al.* Diagnostic accuracy of deep learning in medical imaging: a systematic review
370 and meta-analysis. *npj Digital Medicine* **4**, 65 (2021).
- 371 55. Liu, X. *et al.* A comparison of deep learning performance against health-care professionals in
372 detecting diseases from medical imaging: a systematic review and meta-analysis. *The Lancet Digital*
373 *Health* **1**, e271–e297 (2019).
- 374 56. Nagendran, M. *et al.* Artificial intelligence versus clinicians: systematic review of design, reporting
375 standards, and claims of deep learning studies. *BMJ* vol. 368 m689 (2020).
- 376 57. Riches, N. *et al.* The Effectiveness of Electronic Differential Diagnoses (DDX) Generators: A
377 Systematic Review and Meta-Analysis. *PLoS One* **11**, e0148991 (2016).
- 378 58. Bright, T. J. *et al.* Effect of clinical decision-support systems: a systematic review. *Ann Intern Med*
379 **157**, 29–43 (2012).
- 380 59. Jaspers, M. W. M., Smeulers, M., Vermeulen, H. & Peute, L. W. Effects of clinical decision-support
381 systems on practitioner performance and patient outcomes: a synthesis of high-quality systematic
382 review findings. *J Am Med Inform Assoc* **18**, 327–334 (2011).
- 383 60. Garg, A. X. *et al.* Effects of computerized clinical decision support systems on practitioner
384 performance and patient outcomes: a systematic review. *JAMA* **293**, 1223–1238 (2005).

- 385 61. Varghese, J., Kleine, M., Gessner, S. I., Sandmann, S. & Dugas, M. Effects of computerized decision
386 support system implementations on patient outcomes in inpatient care: a systematic review. *J Am*
387 *Med Inform Assoc* **25**, 593–602 (2018).
- 388 62. Vasey, B. *et al.* Association of Clinician Diagnostic Performance With Machine Learning–Based
389 Decision Support Systems: A Systematic Review. *JAMA Network Open* **4**, e211276–e211276 (2021).
390