

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

Émile Lemoine ^{1,2}, Joel Neves Briard ^{1,3}, Bastien Rioux,^{1,3} Renata Podbielski,³ Bénédicte Nauche,³ Denahin Toffa,^{1,3} Mark Keezer,^{1,4} Frédéric Lesage,² Dang K Nguyen,^{1,3} Elie Bou Assi^{1,3}

To cite: Lemoine É, Neves Briard J, Rioux B, *et al*. Computer-assisted analysis of routine electroencephalogram to identify hidden biomarkers of epilepsy: protocol for a systematic review. *BMJ Open* 2023;**13**:e066932. doi:10.1136/bmjopen-2022-066932

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2022-066932>).

Received 26 July 2022
Accepted 14 January 2023



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Department of Neurosciences, University of Montreal, Montreal, Québec, Canada

²Institute of Biomedical Engineering, Ecole Polytechnique de Montreal, Montreal, Québec, Canada

³University of Montreal Hospital Centre Research Centre, Montreal, Québec, Canada

⁴Stichting Epilepsie Instellingen Nederland (SEIN), Heemstede, The Netherlands

Correspondence to

Dr Émile Lemoine;
emile.lemoine@umontreal.ca

ABSTRACT

Introduction 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 min routine EEGs in patients with epilepsy. In the past three decades, there has been growing interest in the use of computational methods to analyse 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 and analysis 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 analyse 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 modelling, and comment on the reproducibility of the different studies.

Ethics and dissemination Ethical approval was not required. Findings will be disseminated through peer-reviewed publication and presented at conferences related to this field.

PROSPERO registration number CRD42022292261.

BACKGROUND

Epilepsy is characterised by an enduring propensity towards epileptic seizures—transient neurological manifestations provoked

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ This systematic review will be the first to critically evaluate the diagnostic accuracy of computational markers of epilepsy on routine electroencephalogram (EEG), with an emphasis on identifying the barriers towards clinical translation of this technology.
- ⇒ The publication of this protocol ensures transparency, and evaluation of all studies during screening, selection and data extraction by independent reviewers reduces the risk of bias in the selection and analysis of included studies.
- ⇒ High heterogeneity in reporting standards and inclusion criteria is anticipated, possibly preventing the reliable estimation of diagnostic performance metrics.
- ⇒ Our review will constitute a comprehensive reference of current practices in the automated processing and analysis of routine EEG for epilepsy.

by a state of abnormal and excessive neuronal activity in the brain.¹ Epilepsy affects over 65 million people worldwide, and 10% of the population will experience at least one seizure in their lifetime.^{2,3} Epileptic seizures can lead to fractures, road accidents, isolation, anxiety, cognitive decline and death.⁴ In specialised-care settings, the first antiseizure medication (ASM) achieves seizure freedom in approximately 47% of patients.⁵ A prompt diagnosis is key in the prevention of epilepsy-related morbidity and mortality.⁴

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

An EEG records the electrical activity of the brain. It is recommended that all patients who present with a first unprovoked seizure or with new diagnosis of epilepsy undergo an EEG.^{6 7} The initial EEG is generally performed with electrodes applied to the patient's scalp (scalp EEG or routine EEG) for a duration of 20–40 min.⁸ The EEG tracing is then interpreted visually by a neurologist, who attempts to identify interictal epileptiform discharges (IEDs; aka spikes). IEDs are brief (20–200 ms) sharp discharges, clearly emerging from background oscillations, often negative in polarity and sometimes followed by a typical slow wave.⁸ The presence of interictal spikes on the EEG is considered a hallmark of epilepsy, as it represents a strong predictor of seizure recurrence.^{9 10} Furthermore, the identification of interictal spikes can help localise an epileptic focus that may be amenable to surgical resection, and can guide the withdrawal of ASMs in patients after a prolonged period of seizure freedom.^{11 12}

The interictal spike has several limitations. It occurs very sporadically: in patients with epilepsy, only 29%–55% of routine EEGs will capture these transient abnormalities.⁸ After a first unprovoked seizure in adults, the sensitivity of a single routine EEG for detecting epileptiform abnormalities is only 17%.⁹ Furthermore, their identification is somewhat subjective: the percent agreement between EEG experts is around 76%.¹³ Many physiological transient discharges can be misinterpreted as epileptiform spikes. This can lead to the erroneous diagnosis of epilepsy, with sometimes important consequences.^{14 15} In patients labelled with drug-resistant epilepsy, over 25% may have had an erroneous diagnosis as a result of both inadequate history taking and misinterpretation of the EEG.¹⁶ Despite the abundant information on brain activity recorded by the EEG, no other interictal anomalies have been validated for use in clinical settings.^{1 17 18}

Compared with other neuroimaging modalities, a scalp EEG is inexpensive, easy to acquire and confers functional information with high temporal resolution.^{19 20} Moreover, great effort was put in the last decade by the ILAE in standardising the equipment, recording and storage of EEG data.^{10 21} Decades of research have suggested that the automated analysis of EEG can identify hidden differences between with epilepsy and non-epileptic subjects in terms of connectivity,^{22–24} signal predictability and complexity,^{25 26} spectral power^{27 28} and chaoticity.²⁹ Computational analysis of EEG holds the promise of extracting information that is invisible to the naked eye of the human interpreter, in an objective and reproducible manner. Discovering new, non-visible markers of epilepsy could increase the diagnostic yield of the EEG, improve its accessibility and reduce costs, especially in settings where the expertise of a fellowship-trained neurophysiologist is unavailable.^{18 30} In spite of this, none of the proposed non-visible markers of epilepsy have made it into clinical practice.^{10 31} This discrepancy calls attention to the lack of comprehensible and systematic evaluation of these new methods.

We will perform a systematic review of diagnostic test accuracy (DTA) for automated methods of interictal EEG analysis to distinguish between patients with and without epilepsy, without relying on the detection of spikes. The questions that this review addresses are the following: What is the current evidence on the performances of automatically extracted hidden markers compared with the clinical diagnosis of epilepsy by a physician? What is the benefit over the visual identification of IEDs on routine EEG? And what are the different algorithms that have been tested and how does their diagnostic accuracy compare?

METHODS

Study design

This will be a systematic review and meta-analysis following guidance from the Cochrane DTA group. We will report the results according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement for DTA.³²

Study selection criteria

Type of studies

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

Population

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

Reference standard

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

Index test

The index test is a characteristic or feature which is computationally extracted from the EEG signal to identify patients with epilepsy, without relying on detecting IEDs or seizures. These include measures of connectivity, entropy, chaoticity and power spectrum density.³⁵ Also included are statistical models that combine several features or models that take as input the raw or processed EEG.

Search strategy

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

Data items

Data collection will be performed using Covidence by two independent reviewers (EL and JNB/BR), and conflicts will be resolved by a third author (EBA). Authors of the primary study will be contacted if the required data are not available in the original publication. Data collection will include the following information:

1. Title and authors of the study, country of sampling, year of publication.
2. Study type: retrospective versus prospective, design (cohort, case control).
3. Study sample: exclusion and inclusion criteria, number of screened and included patients.
4. Data collection:
 - a. Number of patients, number of EEGs, duration of EEG recordings, use of activation procedures (hyperventilation, photic stimulation, sleep deprivation), setting of recording (hospitalised or ambulatory), whether the same protocol was used for all patients.
 - b. Number of electrodes, sampling frequency.
 - c. If public dataset: reference to the original dataset, dataset name, exclusion/inclusion criteria used on the EEG segments from the dataset.
 - d. Participant characteristics: age, sex, comorbidities, number of ASM, age of first seizure.
5. Reference standard: whether a predefined reference standard was used, definition of reference standard, whether all patients underwent the same reference

standard, time-lapse between reference standard and EEG.

6. Index test:
 - a. Preprocessing: artefact detection and removal (automated or manual), filtering method, filtering frequencies, segmentation protocol (whole EEG vs EEG segments, window size, overlapping vs non-overlapping segments, manual vs automated selection of segments), channel selection.
 - b. Feature extraction and selection: multichannel versus single channel, number of channels selected, whether feature selection was performed, feature extraction algorithm, feature selection method, whether feature selection was applied to data before versus after excluding testing data.
 - c. Classification: algorithm(s) used for classification, testing methodology (cross-validation vs held out testing set).
 - d. Metric used to report diagnostic performances: ROC AUC, accuracy/sensitivity/specificity, F_1 -score, reporting of CIs.
7. Diagnostic performances: number of true positives, number of true negatives, number of false positives, number of false negatives, reported accuracy, reported sensitivity, reported specificity, reported F_1 -score, reported ROC AUC (if more than one index test is performed on the same patient, we will only consider the first test).
8. Reproducibility: whether every data processing step is detailed, whether methods can be reproduced easily, data availability, code availability, open-source computer libraries referenced.

Risk of bias

The risk of bias of all included studies will be assessed through an adapted version of the QUADAS-2 tool.³⁶ Risk of bias for each of the following four elements will be evaluated by two independent reviewers (EL and JNB/BR) as low, high or unclear. Conflicts will be resolved by a third author (EBA). In addition, all publicly available datasets used by at least one of the included studies will be evaluated with the same tool. The following items will be assessed:

1. Patient selection
 - a. Is the population representative of clinical practice?
 - b. Are inclusion and exclusion criteria identical for cases (patients with epilepsy) and controls?
 - c. Are withdrawals explained and appropriate? If individual EEG segments were excluded, were the same criteria used for all segments?
2. Index test
 - a. Were the protocols used for recording the EEG identical in all patients, irrespective of the epilepsy diagnosis?
 - b. Was the index test validated on an independent sample of patients (patients which were not used to identify the index test's threshold or train the learning algorithm)?

3. Reference standard
 - a. Are the criteria used for the diagnosis of epilepsy specified and acceptable (likely to correctly classify the target condition)?
 - b. Was the reference standard assessment independent and blinded to the index test?
4. Flow and timing
 - a. Did the whole sample undergo the reference standard?
 - b. Did the whole sample undergo the same reference standard?
 - c. Was the time-lapse between reference standard and EEG acceptable?
 - d. Was the same data used in the index method available at the time of the reference standard?
 - e. Were all EEGs included in the analysis?

Data synthesis

We will provide a table summarising every published study included in the review, comparing the studies' design, population, reference standard, dataset size, data processing methods and diagnostic accuracy. We will also provide a figure that summarises the risk of bias for each item in the adapted QUADAS-2 tool, comparing (1) every individual article included in the review and (2) every public dataset that is used in ≥ 2 studies.

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

We will summarise in a table the methods used by the different articles during the pipeline's algorithm (preprocessing, feature extraction, feature selection and classification algorithm), along with the proportion of studies that used each method.

Analyses

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

If there are sufficient (≥ 5) studies that report the number of true/false positives and true/false negatives, we will estimate the pooled sensitivity and specificity of each individual marker using a hierarchical, bivariate generalised linear mixed model.³⁷ This allows us to account for the correlation between specificity and sensitivity in a single study. If ≥ 5 studies report these numbers for varying thresholds, we will estimate the pooled ROC curve using the Rutter and Gatsonis hierarchical summary receiver operating characteristic (HSROC) model.³⁸ All analyses will be implemented with the R statistical language. A $p < 0.05$ will be considered statistically significant. Given

insufficient data for the pooled estimates, we will only describe the diagnostic performances (sensitivity, specificity, ROC AUC) narratively. We will present the results of the analyses with forest plots. We will compare the performance of the computational markers for the diagnosis of epilepsy to the visual identification of IEDs on EEG.⁹

We will quantify heterogeneity using the variances of the logit specificity and sensitivity, as well as the median OR.³⁹ The median OR is a measure of inter-study variance translated on the OR scale. It corresponds to the increase in the odds of being true positive/negative in a patient/control going from a study with lower sensitivity/specificity to a study with higher sensitivity/specificity. For heterogeneity in the ROC plane, we will compute the area of the 95% prediction ellipse and present the results on a scatterplot in the ROC plane.³⁹ The median OR and the area of the 95% prediction ellipse are easily obtained and interpreted, and take into account the correlation between a single study's specificity and sensitivity in contrast to univariate methods like Cochrane's *Q* and *I*.^{2 37 40} We will perform subgroup analysis for the following variables: epilepsy type (focal, generalised), epilepsy aetiology (structural vs non-structural), presence of IEDs, age groups (children (<18years), adults (≥ 18 years)), epilepsy syndromes, extracted marker and dataset used. We will also perform a subgroup analysis for populations with a higher prevalence of IEDs without epilepsy (cerebral palsy, autism spectrum disorder, attention deficit disorder)⁴¹ and for extratemporal versus temporal focal epilepsy. We will assess heterogeneity for all subgroup analyses. We will consider a study as belonging to a particular subgroup if $\geq 80\%$ of the studied population belongs to that subgroup. Sensitivity analysis will be conducted for the main analyses by excluding studies with overall high/unclear risk of bias.

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

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

Patient and public involvement

No patients will be involved for this study.

DISCUSSION

The interictal EEG is key in informing the diagnosis of epilepsy, solely based on the visual identification of interictal spikes.⁴² Despite years of research on computational biomarkers of epilepsy, only these spikes are currently used in clinical settings.^{1 17 18} This review aims to systematically evaluate the performances of hidden interictal

markers of epilepsy on EEG against the clinical diagnosis by a physician, describe the data processing pipelines favoured by the researchers to classify the EEG for epilepsy diagnosis and identify the pitfalls that prevent clinical translation of these algorithms.

Algorithms have gained growing interest in medicine for their potential to assist diagnosis and guide clinical decision making.⁴³ EEG analysis is well suited for this application due to the complex nature of the EEG signal. Automated extraction of new epilepsy markers on routine EEG could lead to reduced rate of misdiagnosis, increased availability in areas without access to an expert neurophysiologist and more efficient clinical trials. Research on automatic analysis of EEG data is thriving, in part assisted by the recent increase in computational capacities.^{44–51} However, automatic analysis of EEG is not mentioned in any of the high-quality clinical practice guidelines systematically reviewed by the ILAE.¹⁷

In recent years, increased computational capacities have allowed the development of powerful algorithms that can learn complex representations such as medical images and EEG signals.^{44 52 53} A growing number of algorithms have now been approved by the US Food and Drug Administration for assisting in the diagnosis of several diseases.⁵⁴ Recent systematic reviews have found that most of the studies on automated diagnosis using artificial intelligence have high risk of bias, mostly due to patient selection methodology and absence of validation on external data.^{55–57} Systematic reviews on computer-based clinical-decision support systems also highlight the need for more robust patient selection.^{58–63}

Translation of technology to clinical practice requires strong evidence based on high-quality research. This review is important because it will establish the potential of automatic analysis of EEG as a diagnostic tool for epilepsy, and if evidence to support its use is lacking, it will identify the pitfalls that need to be overcome in future research. Also, by systematically describing current practices that are used by research groups, it will serve as a reference for new researchers in the field. On completion of this review, we will have a better understanding of the potential ways that automated analysis of EEG could be integrated into the clinical workflow; this information will be valuable to anyone designing clinical studies on clinical decision support systems for epilepsy.

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

Twitter Émile Lemoine @lemoineemile, Joel Neves Briard @JNevesBriard, Bastien Rioux @B_Rioux and Elie Bou Assi @eliebouassi

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

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Émile Lemoine <http://orcid.org/0000-0002-8977-3186>

Joel Neves Briard <http://orcid.org/0000-0002-1576-5669>

REFERENCES

- 1 Fisher RS, Acevedo C, Arzimanoglou A, *et al*. ILAE official report: a practical clinical definition of epilepsy. *Epilepsia* 2014;55:475–82.
- 2 Ngugi AK, Bottomley C, Kleinschmidt I, *et al*. Estimation of the burden of active and life-time epilepsy: a meta-analytic approach. *Epilepsia* 2010;51:883–90.
- 3 Hauser WA, Beghi E. First seizure definitions and worldwide incidence and mortality. *Epilepsia* 2008;49:8–12.
- 4 Devinsky O, Spruill T, Thurman D, *et al*. Recognizing and preventing epilepsy-related mortality: a call for action. *Neurology* 2016;86:779–86.
- 5 Kwan P, Brodie MJ. Early identification of refractory epilepsy. *N Engl J Med* 2000;342:314–9.
- 6 Krumholz A, Wiebe S, Gronseth G, *et al*. Practice parameter: evaluating an apparent unprovoked first seizure in adults (an evidence-based review): report of the Quality Standards Subcommittee of the American Academy of Neurology and the American Epilepsy Society. *Neurology* 2007;69:1996–2007.
- 7 Hirtz D, Ashwal S, Berg A, *et al*. Practice parameter: evaluating a first nonfebrile seizure in children: report of the Quality Standards Subcommittee of the American Academy of Neurology, the child Neurology Society, and the American Epilepsy Society. *Neurology* 2000;55:616–23.
- 8 Pillai J, Sperling MR. Interictal EEG and the diagnosis of epilepsy. *Epilepsia* 2006;47:14–22.
- 9 Bouma HK, Labos C, Gore GC, *et al*. The diagnostic accuracy of routine electroencephalography after a first unprovoked seizure. *Eur J Neurol* 2016;23:455–63.
- 10 Tatum WO, Rubboli G, Kaplan PW, *et al*. Clinical utility of EEG in diagnosing and monitoring epilepsy in adults. *Clin Neurophysiol* 2018;129:1056–82.
- 11 Lamberink HJ, Otte WM, Geerts AT, *et al*. Individualised prediction model of seizure recurrence and long-term outcomes after withdrawal of antiepileptic drugs in seizure-free patients: a systematic review and individual participant data meta-analysis. *Lancet Neurol* 2017;16:523–31.
- 12 West S, Nevitt SJ, Cotton J, *et al*. Surgery for epilepsy. *Cochrane Database Syst Rev* 2019;6:CD010541.
- 13 Jing J, Herlopian A, Karakis I, *et al*. Interrater reliability of experts in identifying interictal epileptiform discharges in electroencephalograms. *JAMA Neurol* 2020;77:49–57.

- 14 Amin U, Benbadis SR. The role of EEG in the erroneous diagnosis of epilepsy. *J Clin Neurophysiol* 2019;36:294–7.
- 15 Kang JY, Krauss GL. Normal variants are commonly overread as interictal epileptiform abnormalities. *J Clin Neurophysiol* 2019;36:257–63.
- 16 Smith D, Defalla BA, Chadwick DW. The misdiagnosis of epilepsy and the management of refractory epilepsy in a specialist clinic. *QJM* 1999;92:15–23.
- 17 Sauro KM, Wiebe S, Dunkley C, et al. The current state of epilepsy guidelines: a systematic review. *Epilepsia* 2016;57:13–23.
- 18 Engel J, Bragin A, Staba R. Nonictal EEG biomarkers for diagnosis and treatment. *Epilepsia Open* 2018;3:120–6.
- 19 DellaBadia J, Bell WL, Keyes JW, et al. Assessment and cost comparison of sleep-deprived EEG, MRI and PET in the prediction of surgical treatment for epilepsy. *Seizure* 2002;11:303–9.
- 20 Abdelhady S, Shokri H, Fathy M, et al. Evaluation of the direct costs of epilepsy in a sample of Egyptian patients following up in Ain Shams university hospital. *Egypt J Neurol Psychiatry Neurosurg* 2020;56:112.
- 21 Velis D, Plouin P, Gotman J, et al. Recommendations regarding the requirements and applications for long-term recordings in epilepsy. *Epilepsia* 2007;48:379–84.
- 22 Schmidt H, Woldman W, Goodfellow M, et al. A computational biomarker of idiopathic generalized epilepsy from resting state EEG. *Epilepsia* 2016;57:e200–4.
- 23 Lopes MA, Perani S, Yaakub SN, et al. Revealing epilepsy type using a computational analysis of interictal EEG. *Sci Rep* 2019;9:10169.
- 24 Verhoeven T, Coito A, Plomp G, et al. Automated diagnosis of temporal lobe epilepsy in the absence of interictal spikes. *Neuroimage Clin* 2018;17:10–5.
- 25 Ouyang C-S, Yang R-C, Wu R-C, et al. Determination of antiepileptic drugs withdrawal through EEG Hjorth parameter analysis. *Int J Neural Syst* 2020;30:2050036.
- 26 Zhang J-H, Han X, Zhao H-W, et al. Personalized prediction model for seizure-free epilepsy with levetiracetam therapy: a retrospective data analysis using support vector machine. *Br J Clin Pharmacol* 2018;84:2615–24.
- 27 Oliva JT, Rosa JLG. Differentiation between normal and interictal EEG using multitaper spectral classifiers. 2018 international joint conference on neural networks (IJCNN). *Rio de Janeiro* 2018;1–8.
- 28 Pegg EJ, Taylor JR, Mohanraj R. Spectral power of interictal EEG in the diagnosis and prognosis of idiopathic generalized epilepsies. *Epilepsy Behav* 2020;112:107427.
- 29 Jacob JE, Sreelatha VV, Iype T, et al. Diagnosis of epilepsy from interictal EEGs based on chaotic and wavelet transformation. *Analog Integr Circ Sig Process* 2016;89:131–8.
- 30 Wahl B, Cossy-Gantner A, Germann S, et al. Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? *BMJ Glob Health* 2018;3:e000798.
- 31 Pitkänen A, Löscher W, Vezzani A, et al. Advances in the development of biomarkers for epilepsy. *Lancet Neurol* 2016;15:843–56.
- 32 McInnes MDF, Moher D, Thoms BD, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *JAMA* 2018;319:388–96.
- 33 November J. Biomedical computing: digitizing life in the United States. In: *Biomedical computing: digitizing life in the United States*. 2012: 1–344.
- 34 Fisher RS, van Emde Boas W, Blume W, et al. Epileptic seizures and epilepsy: definitions proposed by the International League against epilepsy (ILAE) and the International bureau for epilepsy (IBE). *Epilepsia* 2005;46:470–2.
- 35 Supriya S, Siuly S, Wang H, et al. Automated epilepsy detection techniques from electroencephalogram signals: a review study. *Health Inf Sci Syst* 2020;8:33.
- 36 Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529–36.
- 37 Reitsma JB, Glas AS, Rutjes AWS, et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;58:982–90.
- 38 Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 2001;20:2865–84.
- 39 Plana MN, Pérez T, Zamora J. New measures improved the reporting of heterogeneity in diagnostic test accuracy reviews: a metaepidemiological study. *J Clin Epidemiol* 2021;131:101–12.
- 40 Rucker G, Schwarzer G, Carpenter JR, et al. Undue reliance on I (2) in assessing heterogeneity may mislead. *BMC Med Res Methodol* 2008;8:79.
- 41 So EL. Interictal epileptiform discharges in persons without a history of seizures: what do they mean? *J Clin Neurophysiol* 2010;27:229–38.
- 42 Smith SJM. EEG in the diagnosis, classification, and management of patients with epilepsy. *J Neurol Neurosurg Psychiatry* 2005;76:ii2–7.
- 43 Obermeyer Z, Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. *N Engl J Med* 2016;375:1216–9.
- 44 Roy Y, Banville H, Albuquerque I, et al. Deep learning-based electroencephalography analysis: a systematic review. *J Neural Eng* 2019;16:051001.
- 45 Craik A, He Y, Contreras-Vidal JL. Deep learning for electroencephalogram (EEG) classification tasks: a review. *J Neural Eng* 2019;16:031001.
- 46 Rasheed K, Qayyum A, Qadir J, et al. Machine learning for predicting epileptic seizures using EEG signals: a review. *IEEE Rev Biomed Eng* 2021;14:139–55.
- 47 Gemein LAW, Schirrmeyer RT, Chrabaszcz P, et al. Machine-learning-based diagnostics of EEG pathology. *Neuroimage* 2020;220:S1053–8119(20)30507–3.
- 48 Mesraoua B, Deleu D, Al Hail H, et al. Electroencephalography in epilepsy: look for what could be beyond the visual inspection. *Neuro Sci* 2019;40:2287–91.
- 49 van Dieën E, Zweiphenning WJEM, Jansen FE, et al. Brain network organization in focal epilepsy: a systematic review and meta-analysis. *PLOS ONE* 2014;9:e114606.
- 50 Faiman I, Smith S, Hodsoll J, et al. Resting-state EEG for the diagnosis of idiopathic epilepsy and psychogenic nonepileptic seizures: a systematic review. *Epilepsy Behav* 2021;121:S1525–5050(21)00281–X.
- 51 Pegg EJ, Taylor JR, Keller SS, et al. Interictal structural and functional connectivity in idiopathic generalized epilepsy: a systematic review of graph theoretical studies. *Epilepsy Behav* 2020;106:S1525–5050(20)30192–X.
- 52 Esteva A, Robicquet A, Ramsundar B, et al. A guide to deep learning in healthcare. *Nat Med* 2019;25:24–9.
- 53 Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal* 2017;42:60–88.
- 54 American College of Radiology Data Science Institute. FDA cleared AI algorithms. n.d. Available: <https://models.acrdsi.org>
- 55 Aggarwal R, Sounderajah V, Martin G, et al. Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. *NPJ Digit Med* 2021;4:65.
- 56 Liu X, Faes L, Kale AU, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: A systematic review and meta-analysis. *Lancet Digit Health* 2019;1:e271–97.
- 57 Nagendran M, Chen Y, Lovejoy CA, et al. Artificial intelligence versus clinicians: systematic review of design, reporting Standards, and claims of deep learning studies. *BMJ* 2020;368:m689.
- 58 Riches N, Panagioti M, Alam R, et al. The effectiveness of electronic differential diagnoses (DDX) generators: a systematic review and meta-analysis. *PLoS One* 2016;11:e0148991.
- 59 Bright TJ, Wong A, Dhurjati R, et al. Effect of clinical decision-support systems: a systematic review. *Ann Intern Med* 2012;157:29–43.
- 60 Jaspers MWM, Smeulers M, Vermeulen H, et al. Effects of clinical decision-support systems on practitioner performance and patient outcomes: a synthesis of high-quality systematic review findings. *J Am Med Inform Assoc* 2011;18:327–34.
- 61 Garg AX, Adhikari NKJ, McDonald H, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *JAMA* 2005;293:1223–38.
- 62 Varghese J, Kleine M, Gessner SI, et al. Effects of computerized decision support system implementations on patient outcomes in inpatient care: a systematic review. *J Am Med Inform Assoc* 2018;25:593–602.
- 63 Vasey B, Ursprung S, Beddoe B, et al. Association of clinician diagnostic performance with machine learning-based decision support systems: a systematic review. *JAMA Netw Open* 2021;4:e211276.