Development and validation of a diagnostic aid for convulsive epilepsy in sub-Saharan Africa: a retrospective case-control study

*Gabriel Davis Jones, Symon M Kariuki, Anthony K Ngugi, Angelina Kakooza Mwesige, Honorati Masanja, Seth Owusu-Agyei, Ryan Wagner, J Helen Cross, Josemir W Sander, Charles R Newton, Arjune Sen, on behalf of the EPInA Study Group**

Summary

Background Identification of convulsive epilepsy in sub-Saharan Africa relies on access to resources that are often unavailable. Infrastructure and resource requirements can further complicate case verification. Using machine-learning techniques, we have developed and tested a region-specific questionnaire panel and predictive model to identify people who have had a convulsive seizure. These findings have been implemented into a free app for health-care workers in Kenya, Uganda, Ghana, Tanzania, and South Africa.

Methods In this retrospective case-control study, we used data from the Studies of the Epidemiology of Epilepsy in Demographic Sites in Kenya, Uganda, Ghana, Tanzania, and South Africa. We randomly split these individuals using a 7:3 ratio into a training dataset and a validation dataset. We used information gain and correlation-based feature selection to identify eight binary features to predict convulsive seizures. We then assessed several machine-learning algorithms to create a multivariate prediction model. We validated the best-performing model with the internal dataset and a prospectively collected external-validation dataset. We additionally evaluated a leave-one-site-out model (LOSO), in which the model was trained on data from all sites except one that, in turn, formed the validation dataset. We used these features to develop a questionnaire-based predictive panel that we implemented into a multilingual app (the Epilepsy Diagnostic Companion) for health-care workers in each geographical region.

Findings We analysed epilepsy-specific data from 4097 people, of whom 1985 (48·5%) had convulsive epilepsy, and 2112 were controls. From 170 clinical variables, we initially identified 20 candidate predictor features. Eight features were removed, six because of negligible information gain and two following review by a panel of qualified neurologists. Correlation-based feature selection identified eight variables that demonstrated predictive value; all were associated with an increased risk of an epileptic convulsion except one. The logistic regression, support vector, and naive Bayes models performed similarly, outperforming the decision-tree model. We chose the logistic regression model for its interpretability and implementability. The area under the receiver operator curve (AUC) was 0·92 (95% CI 0·91–0·94, sensitivity 85·0%, specificity 93·7%) in the internal-validation dataset and 0·95 (0·92–0·98, sensitivity 97·5%, specificity 82·4%) in the external-validation dataset. Similar results were observed for the LOSO model (AUC 0·94, 0·93–0·96, sensitivity 88·2%, specificity 95·3%).

Interpretation On the basis of these findings, we developed the Epilepsy Diagnostic Companion as a predictive model and app offering a validated culture-specific and region-specific solution to confirm the diagnosis of a convulsive epileptic seizure in people with suspected epilepsy. The questionnaire panel is simple and accessible for health-care workers without specialist knowledge to administer. This tool can be iteratively updated and could lead to earlier, more accurate diagnosis of seizures and improve care for people with epilepsy.

Funding The Wellcome Trust, the UK National Institute of Health Research, and the Oxford NIHR Biomedical Research Centre.

Introduction

Epilepsy affects more than 50 million people worldwide and contributes to 0·5% of the global disease burden.[1] Much of this load is in low-income and middle-income countries (LMICs), at least in part attributable to epilepsy risk factors.[2,3] Most LMIC data are based on convulsive epilepsy, namely tonic-clonic seizures. Convulsive epileptic seizures are more easily recognised and associated with more substantial stigma and mortality than non-convulsive seizures.

Diagnosing epilepsy is difficult, particularly in LMICs. Few neurologists or other health-care professionals have the necessary expertise and time to acquire the detailed history needed. Reliable and robust technologies facilitating a diagnosis would therefore be helpful. Apps on smartphones and tablets have been shown to aid the

Lancet Digit Health 2023; 5: e185–93

*Members of the EPInA Study Group are listed in the appendix

Oxford Epilepsy Research Group, NIHR Biomedical Research Centre, Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital, Oxford, UK (G D Jones MD, Prof A Sen PhD); **The Alan Turing Institute, London, UK** (G D Jones); **KEMRI and Wellcome Trust Research Programme, Centre for Geographic Medicine Research Coast, Kilifi, Kenya** (S M Kariuki PhD, Prof C R Newton MD); **Department of Population Health** (A K Ngugi PhD) **and Brain and Mind Institute** (A K Ngugi), **Aga Khan University, Nairobi, Kenya; Centre for Global Health Equity, University of Michigan, Ann Arbor, MI, USA** (A K Ngugi); **Studies of Epidemiology of Epilepsy in Demographic Surveillance Systems, INDEPTH Network, Accra, Ghana** (S M Kariuki, Prof C R Newton); **Department of Paediatrics and Child Health, Makerere University College of Health Sciences, Kampala, Uganda** (A K Mwesige PhD); **Ifakara Health Institute, Ifakara, Tanzania** (H Masanja PhD); **Kintampo Health Research Centre, Kintampo, Ghana** (Prof S Owusu-Agyei PhD); **MRC and Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa** (R Wagner PhD); **Developmental Neurosciences, University College London NIHR BRC Great Ormond Street Institute of Child Health, London, UK** (Prof J H Cross MD); **Department of Clinical and Experimental Epilepsy, UCL Queen Square**

Institute of Neurology, London, UK (Prof J W Sander PhD); **Stichting Epilepsie Instellingen Nederland, Heemstede, Netherlands** (Prof J W Sander); **Department of Psychiatry, University of Oxford, Oxford, UK** (Prof C R Newton)

Correspondence to: Prof Arjune Sen, Oxford Epilepsy Research Group, NIHR Biomedical Research Centre, Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital, Oxford OX3 9DU, UK
arjune.sen@ndcn.ox.ac.uk

See Online for appendix

Research in context

Evidence before this study

We searched PubMed for publications in English from inception until Sept 1, 2022, for predictive models to identify convulsive seizures in sub-Saharan Africa, using the terms "convulsive epilepsy", "seizure", "prediction model", "app", "low and middle income coutries (LMIC)", and "sub-Saharan Africa". Previous studies have reported models using several risk factors to predict convulsive epilepsy. These models have either not established data-derived risk factors, not used regional data, have comprised small cohorts, or have underexplored potential predictive algorithms.

Added value of this study

We used robust machine-learning methods to develop and test a reliable predictive algorithm for the presence of convulsive epilepsy. We used a large epilepsy-specific dataset from five regions across sub-Saharan Africa, comprising more than 4000 people. We then prospectively validated the predictive algorithm in Kilifi, Kenya. The algorithm, incorporating eight binary questions, was implemented into a free, multilingual app for health-care workers, which will now be tested prospectively in other regions.

Implications of all the available evidence

A region and culture-specific diagnostic aid can allow for reliable and rapid identification of people with convulsive epileptic seizures. This rapid identification would potentially reduce the diagnostic gap and improve individual risk mitigation and streamline the care pathway. Further, this new tool can reduce the burden on health-care resources in low-income and middle-income countries by empowering community health-care workers.

diagnosis of epilepsy in primary-care settings in LMICs, particularly in busy clinics.[4] It is essential that any such technologies are culturally tailored and derived from data acquired from local populations. We developed a predictive model and app to help diagnose convulsive epilepsies in African populations on the basis of data collected across Africa.

Methods

Study design and participants

In this retrospective case-control study, we analysed data from the Studies of the Epidemiology of Epilepsy in Demographic Sites (SEEDS) database. The protocol for SEEDS is published elsewhere.[5] SEEDS was carried out in five Health Demographics Surveillance System (HDSS) sites, which are part of the International Network for the Demographic Evaluation of Populations and Their Health in Low and Middle-Income Countries. We analysed data from surveys conducted in the following areas: Kilifi, Kenya (carried out between Dec 3, 2007, and July 31, 2008); Iganga-Mayuge, Uganda (carried out between Feb 2, 2009, and Oct 30, 2009); Ifakara, Tanzania (carried out between May 4, 2009, and Dec 31, 2009); Agincourt, South Africa (carried out between Aug 4, 2008, and Feb 27, 2009); and Kintampo, Ghana (carried out between Aug 2, 2010, and April 29, 2011).

A three-stage screening process was used to identify cases.[6] The first stage involved applying two screening questions during the routine, door-to-door census in each HDSS centre. Heads of households were interviewed about convulsions in each individual at the dwelling. In the second stage, a more detailed questionnaire was administered by field workers to individuals who screened positive in the first stage.[7] In the third stage, people positive at the second stage were assessed by clinicians who made a final diagnosis of epilepsy. This data-acquisition process was identical across all sites.

We only evaluated active convulsive epilepsies, as convulsions are more reliably detected, more likely to be reported, and associated with a greater mortality risk than non-convulsive seizures.[8] Active convulsive epilepsy was defined as two or more unprovoked convulsions (which could be generalised at onset or focal to bilateral tonic-clonic seizures) occurring at least 24 h apart with at least one episode in the preceding 12 months. For each case of active convulsive epilepsy identified in the third stage, a community control matched by age and sex was randomly selected from the census database of the relevant centre for a case-control study. Fieldworkers administered questionnaires on the basis of previous studies[9,10] to cases and controls and were masked to the status of the individual. Sociodemographic variables and historical risk factors (perinatal events, head injuries, and diet) were collected. The medical history was obtained, and a diagnosis was made by trained clinicians. The parent or caregiver was interviewed if a participant was younger than 18 years or had cognitive impairment.

In our study, the individuals included were identified within the census as having positive indicators of convulsive epilepsy (a history of convulsions in a lifetime). Each participant was phenotyped according to approximately 170 unique variables (features) in the following categories: clinical history; clinical examination; seizure description; and electroencephalogram (EEG) interpretation. The primary outcome was a diagnosis of convulsive epilepsy with an EEG-aided diagnosis confirmed by a neurologist specialised in epilepsy. Controls were people without a diagnosis of convulsive epilepsy who completed the same clinical pathway. We removed outliers and erroneous entries and replaced these with missing values. Children younger than six years or people without a definitive diagnosis were excluded.

We aimed to develop a maximally discriminative predictive model relying solely on binary information for

ease of use by health-care workers and to reduce the possibility of data-entry errors. We therefore removed any phenotypic features not easily reproduced as binary questions. These features included physical examination, seizure subtype information, and EEG findings. Continuous or categorical variables that could not be discretised, binarised, or that were missing more than 40% of their information were excluded.

Data processing, feature selection, and machine-learning algorithms

We examined the entire dataset for missing values and imputed missing information using stochastic regression (appendix p 1). Clinical variables for which 100% of the data were available were used to develop the imputation model. An adaptive multivariate standard approximation rounding procedure was then used to binarise the imputed data points. Data were then spliced (7:3) into model training and internal validation subsets balanced in the outcome class, geographical region, and sex (appendix p 5). We chose this ratio to enable robust model development from the training dataset while retaining enough data for model validation (figure 1). Although the SEEDS dataset was a case-control, randomly matched study that used age and sex, we chose not to undo the matching procedure, given that age and sex are not generally regarded as potential confounders in the clinical diagnosis of epilepsy. The diagnosis of convulsive epilepsy is made primarily on the basis of clinical history and what the person experiences before, during, and after the convulsive episode.

We used two established feature analysis and selection methods to identify suitable features, which were information gain and correlation-based feature selection. Information gain measures how much the knowledge of one feature reduces the uncertainty of the outcome.[11] In keeping with similar studies, we used a threshold of 0·01 to establish which features possessed negligible information gain.[12] Correlation-based feature selection (CFS) is a multivariable filter that explores intrafeature relationships, independent of the outcome class.[13] Features with a high correlation with at least one other feature are more linearly dependent (multicollinear) and will share a similar relationship with the predicted class. Multicollinear variables can be considered redundant for predictive modelling and all except one variable (typically the variable with the strongest correlation with the outcome class) can be removed. Incorporating multicollinear features in a model results in redundant features and inaccurate feature-weighting estimates, weakening the statistical power of the model.[14] For our study, an appropriate subset of features would contain those highly correlated with convulsive epilepsy, yet minimally correlated with each other.

Information gain and CFS are selection methods for filtering candidate features. These methods are also independent of the predictive algorithm, enabling improved interpretability of selected features and their clinical relevance while minimising model overfitting. This is in contrast to other feature-selection methods, such as stepwise logistical regression or backward-elimination regression.

Adding excessive numbers of features to a predictive model decreases performance. With an increase in the

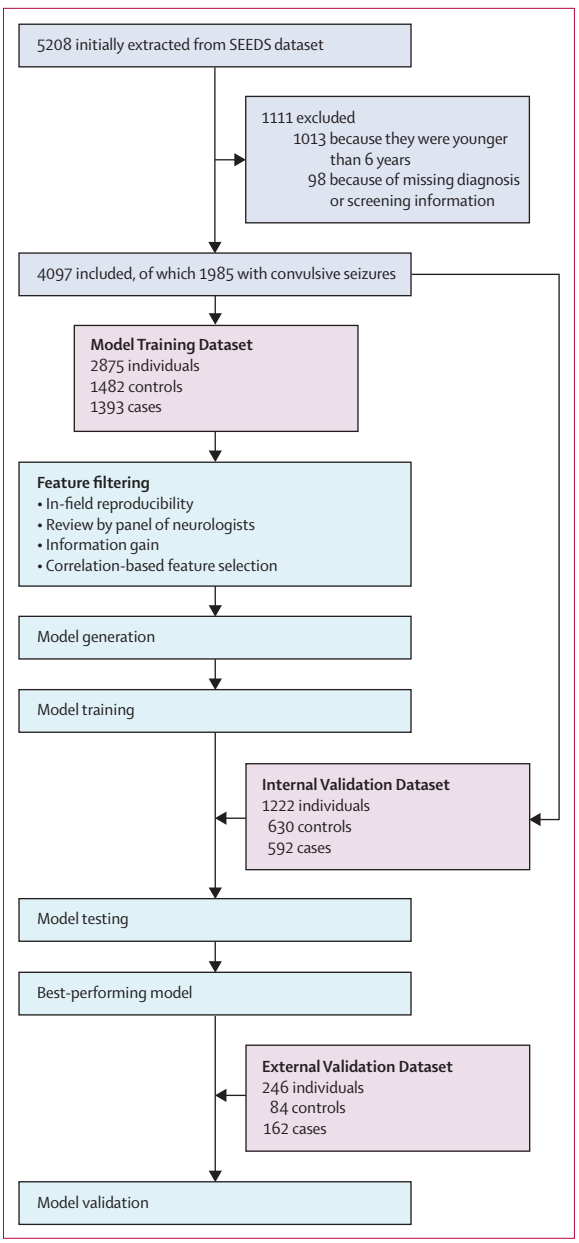

Figure 1: **Workflow for data processing and model development**
The SEEDs dataset from five unique regions in sub-Saharan Africa (Ghana, Kenya, South Africa, Tanzania, and Uganda) was spliced into two subsets balanced for prevalence of convulsive epilepsy, sex, and geographical site, with n=2875 for model training and n=1222 for validation. The best-performing model was then selected and evaluated using the prospectively collected external-validation dataset from Kilifi, Kenya.

number of features, the risk of model overfitting increases. To identify the optimum number of features required, we iteratively identified k features that showed a correlation with convulsive epilepsy and plotted the change in mean area under the receiver operator curve (AUC) to predict convulsive epilepsy using between 1 and k features, in which k is a previously unknown quantity. This approach results in more robust models while avoiding overfitting. Local board-certified neurologists with region-specific knowledge then reviewed the selected features to remove questions that were confounders even when representing a strong correlation with the outcome class.

We then trained four machine-learning algorithms, comprising a decision tree, logistic regression, naive Bayes (assuming a Bernoulli distribution), and a support-vector machine (using a linear-bias kernel). We chose these models because their results are easily interpretable, can be used without substantial computational requirements, and are established for use in clinically relevant predictive applications.[12,15,16]

We developed the predictive models using ten-fold cross validation; 90% of the training dataset was used to train the model and the remaining 10% of the training dataset was used to evaluate the performance of the model. This cross-validation was then repeated a further nine times (ten folds in total), with each fold using a new 90% of the training dataset and 10% for evaluation. Using the AUC from each fold, we compared the mean AUC between each model using the Kruskall-Wallis one-way ANOVA. We then ranked the models in descending order of mean AUC and excluded significantly poorer models identified using a pair-wise Mann-Whitney U test with a significance threshold of 0·01.

Then, we tested the best-performing model on the internal validation dataset. We used model intercept correction to adjust the model for a post-screening prevalence of 23·8% on the basis of previous estimates in the literature.[5] The same prevalence was used in subsequent analyses. To determine the probability threshold for a classification of probable epilepsy, we identified the threshold corresponding to a minimum sensitivity and specificity of 80%.

We evaluated the sensitivity, specificity, positive predictive value, and negative predictive value. To assess differences in clinical sites, we repeated this process, using four clinical sites for model training and one for model testing (leave one site out [LOSO]) to evaluate inter-site consistency.

We then performed a decision-curve analysis using the internal validation dataset to compare the net benefit of the best-performing model against a treat-all strategy.[17] In this context, a treat-all approach would be one in which people with suspected epilepsy who passed the screening criteria would be either directly prescribed an anti-seizure medication or immediately referred to a clinic instead of being assessed with the predictive model. A robust model should outperform the treat-all strategy across a range of probability thresholds to identify people needing treatment while avoiding unnecessary intervention in people unlikely to have convulsive epilepsy. An ineffective model would show either no net benefit or a lower net benefit relative to the treat-all strategy. We evaluated probability thresholds between 0·01 and 0·99 (1–99%). We additionally assessed the calibration degree using the Brier score.

External validation with pilot data

To mitigate potential changes in the diagnosis of convulsive epilepsy since the model-development data were first acquired, we evaluated the performance of our model using prospectively-collected pilot data. We acquired validation data from the KEMRI-Wellcome Health Research Unit in Kilifi, Kenya, between Nov 15, 2021, and March 7, 2022. People older than 6 years of age with suspected or confirmed epilepsy attending the local clinic were first screened using the same criteria outlined in the model development. Participants with confirmed convulsive epilepsy were used as cases whereas those without a diagnosis of convulsive epilepsy were controls.

Development of a smartphone app

Using the feature weights of the final model, we developed an app, the Epilepsy Diagnostic Companion (EDC), for use by community-based health-care workers. The EDC is an Android-based smartphone app developed using Java JDK 11.0.7 within Android Studio version 4.0 and is designed to run on Android Application Programming Interface 16 or higher, enabling compatibility with 99·8% of Android devices. We chose to develop for Android devices, as these are more common than other operating systems in African nations (Android operating system prevalence 83·9% *vs* Apple operating system 13·2%, as of September, 2022).[18] The screening questions and selected features were converted into a binary (yes or no) questionnaire and the responses were used to provide the probability and classification (likely or unlikely) of convulsive epilepsy. An accompanying clinical report comprising anonymised metadata and questionnaire responses is provided. The app and clinical questions were also translated into the languages represented at each of the study sites.

Users can export the anonymised report for the person with suspected convulsive epilepsy, and data are also stored on a cloud-based, encrypted server. For each report, the server stores associated metadata (eg, timestamp and coarse GPS location), the questionnaire responses, and the prediction of the model. A unique one-word key is also assigned to the report to allow for three-factor identification of the individual by a prenominated recipient clinic. The individual can receive the timestamp, location, and unique key for their report via SMS or email. These three independently anonymous identifiers

can then be used to link the person with their report in the clinic without compromising their identity.

The potential biases of this study were as follows: there was high attrition between survey stages because of logistical difficulties in contacting eligible follow-up participants; prevalence estimates from these studies may not reflect other regions of sub-Saharan Africa given that the study sites were selected on the criteria of endemicity of potential risk factors and the availability of minimum resources required to support the studies; effect estimates from this study require cautious interpretation, particularly those that are susceptible to recall bias (eg, perinatal risk factors) and measurement error (eg, under-reporting of diabetes); selection bias, meaning the screening criteria implemented in the SEEDS study may differ from those implemented in other environments (eg, if people with suspected epilepsy are screened with additional screening criteria from those implemented in this study, then the observed performance of the predictive model may differ); and prevalence bias, meaning if the underlying prevalence of convulsive epilepsy is different in a specific population, then the positive and negative predictive values of the predictive model may differ.

Statistical analysis

We adhered to TRIPOD guidelines for reporting.[19] All available data were used in feature analysis, model training, testing, and validation. Discrete variables were presented as numbers and percentages whereas continuous variables were listed as mean and SD. Candidate predictive models were compared using the Kruskal-Wallis one-way ANOVA and pair-wise Mann-Whitney U test with a significance threshold of 0·01. p values were estimated for each feature's association with the diagnosis of convulsive epilepsy using the Mann-Whitney U test and a significance threshold of 0·05. The risk of convulsive epilepsy associated with each feature was estimated using odds ratios and 95% CIs. The CIs for sensitivity and specificity were the exact Clopper-Pearson intervals and those for predictive values were the standard Mercaldo logit CIs. Model intercept adjustment was applied using the method described by Huang and colleagues.[20] Analysis was done using R (version 3.6.0) in RStudio (version 1.2.5019).

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

The demographic and symptom characteristics of the SEEDs study have been published elsewhere.[5] 5208 individuals were initially extracted from the SEEDs database. Individuals younger than 6 years of age (n=1013) and people missing a diagnosis (n=84) were excluded. 14

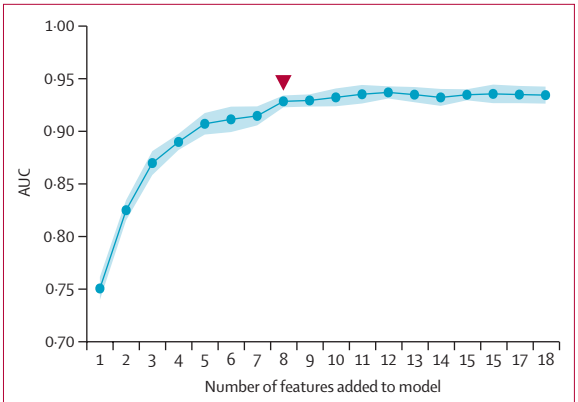

Figure 2: **Identification of optimal number of predictive features for a prediction of convulsive epilepsy**
Increasing the number of features strengthens the model to a plateau point around eight features (red arrow). The model AUC does not improve significantly beyond eight features (rolling mean p=0·77), even when up to 18 features are added (AUC with eight features 0·93). This demonstrates that adding features does not necessarily result in improved model performance (AUC with 18 features 0·93). Blue indicates 95% CI. AUC=area under the receiver operator curve.

	Information gain	Selected by correlation-based feature selection	Regression coefficients	Odds ratio for convulsive epilepsy (95% CI)
During these episodes have you ever bitten your tongue?	0·259	Yes	2·10	8·14 (5·51–12·23)
Have you ever wet yourself during these episodes?	0·251	Yes	3·00	20·11 (13·18–31·56)
During these episodes, do you lose contact with your surroundings?	0·139	Yes	1·66	5·25 (3·37–8·29)
Has anyone told you that you appear dazed during these episodes?	0·128	Yes	1·82	6·18 (3·92–9·89)
During these episodes does your body stiffen?	0·117	Yes	2·29	9·87 (6·32–15·81)
Do you frequently not remember these episodes or do you ever find yourself in a place or position and you do not know how you got there?	0·099	No	..	..
Have you ever been told that your arms, legs, or body twitch or jerk during these episodes?	0·087	No	..	..
Do you experience stomach ache before these episodes?	0·086	Yes	2·13	8·43 (4·39–17·19)
Do you see odd things (eg, flashes or bright lights) before these episodes occur?	0·081	Yes	1·81	6·11 (2·85–14·05)
Do you think anything brings on these episodes?	0·056	Yes	−3·04	0·05 (0·03–0·07)
Do your arms, legs or face shake or tremble during these episodes?	0·045	No	..	..
Do you experience odd smells before these episodes?	0·025	No	..	..

These features demonstrated non-negligible information gain (>0·01) to predict a diagnosis of convulsive epilepsy and were ranked in descending order. The number of features were then filtered using correlation-based feature selection. The final eight features were used to develop the predictive model. The adjusted intercept for the logistic regression equation is −2·58 (95% CI −2·65 to −2·51).

Table 1: **Candidate features for a classification of convulsive epilepsy**

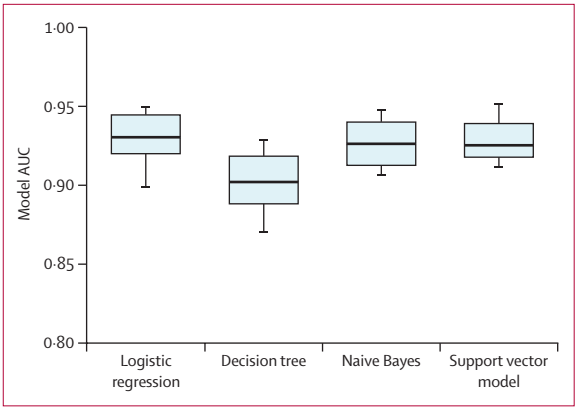

Figure 3: **Comparison of model AUC with different machine learning algorithms**
Central line indicates mean AUC, lower edge of the box indicates first quartile, upper edge of the box indicates third quartile, lower whisker indicates minimum AUC, and upper whisker indicates maximum AUC. Decision tree, logistic regression, naive Bayes assuming a Bernoulli distribution, and a support-vector machine using a linear kernel are shown. The logistic regression, naive Bayes, and support vector models showed similar results (p=0·67), and outperformed the decision-tree model (p<0·001). AUC=receiver operator characteristic area under the curve.

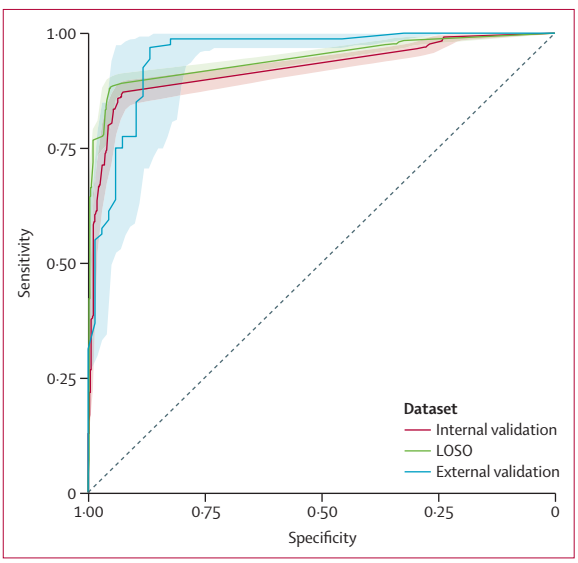

Figure 4: **AUC for the logistic regression model trained to predict convulsive epilepsy**
The internal-validation dataset AUC was 0·92 (95% CI 0·91–0·94), the AUC was 0·94 (0·93–0·96), and the external validation dataset AUC was 0·96 (0·93–0·98). AUC=receiver operator characteristic area under the curve. LOSO=leave one site out.

individuals with incomplete screening information were excluded. Of the 4097 people with a suspected convulsive seizure, 1985 (48·5%) were diagnosed with convulsive epilepsy. This number included those with a previous diagnosis of the condition (figure 1). The median age of the sample was 19 years (IQR 12·3–31·9). 2158 (52·7%) individuals were female and 1939 (47·3%) were male. The proportion of people with epileptic convulsive seizures was higher in males (986 [50·9%] of 1939) than in females (999 [46·3%] of 2158).

The proportions of individuals with convulsive epilepsy following preprocessing and screening were 310 (48·3%) of 642 in Agincourt, 413 (42·0%) of 984 in Ifakara, 197 (43·2%) of 456 in Iganga, 687 (54·6%) of 1259 in Kilifi, and 378 (50·0%) of 756 in Kintampo (appendix p 4). We used an estimated prevalence after screening of 23·8% to assess model performance.[5]

We assessed for features with non-negligible information gain (>0·01) that would positively affect a predictive model of convulsive epilepsy from the model training dataset. 20 potential features were identified. Six demonstrated negligible information gain and two were removed following review by a panel of certified neurologists. These two features related to whether the individual was concurrently taking any medications or traditional medicines. The sex of the person with suspected epilepsy was not identified as a significant predictor (information gain <0·001, p=0·52).

We ranked the remaining 18 features from highest to lowest information gain. We evaluated subsets with the top k features ranging between one and 18 selected features to identify the threshold at which adding features to the predictive model would not significantly improve model performance. We identified eight features with the highest information gain and found no significant increase in the AUC after including these features (AUC 0·93, p=0·69; figure 2). This finding suggested that adding more features, even those strongly correlated with convulsive seizures, would not necessarily improve model performance (mean rolling p value for the remaining feature sets 0·77). We then applied CFS on the initial 18 features and identified eight features as independent predictors of convulsive epilepsy. These features were not the same as the top eight features identified with information gain (table 1).

We trained four machine-learning algorithms using the features identified following CFS. The logistic regression, support vector machine, and naive Bayes models performed similarly (mean AUC for logistic regression 0·93, 95% CI 0·92–0·94; mean AUC for support vector model 0·93, 0·92–0·94; and mean AUC for naive Bayes AUC 0·93, 0·92–0·94; p=0·67). These models outperformed the decision tree (mean AUC 0·90, 0·89–0·92; p<0·01; figure 3). We selected the logistic regression model for further analysis because its coefficients were generally the simplest to interpret and easiest to implement into a clinical tool using digital or manual calculation. We then retrained the logistic regression model on the complete model training dataset and assessed its performance using the internal validation dataset and an adjusted model intercept using a prevalence of 23·8%.

Decision-curve analysis demonstrated the net benefit of the model exceeded a treat-all strategy for all probability thresholds of 0·11 or higher (appendix p 2). Below this

threshold the absolute average difference in the net benefit between the model and the treat-all strategy was 0·0086. The calibration degree (Brier score) for the model was 0·11, indicating a high degree of calibration. When tested on the internal validation dataset, the model performed well with an AUC of 0·92 (95% CI 0·91–0·94; figure 4). We selected a probability threshold of 0·28 (28%) which corresponded with a sensitivity of 85·0% (95% CI 81·8–87·7), specificity of 93·7% (91·5–95·4), positive-predictive value of 92·6% (90·1–94·7), and negative predictive value of 86·9% (84·1–89·3) in the internal validation dataset (table 2).

Similar results were observed for the LOSO model. For predicting convulsive epilepsy in the cohort from Kilifi when the model had not been trained using data from this site, the AUC was 0·94 (95% CI 0·93–0·96), sensitivity 88·2% (85·6–90·5), specificity 95·3% (93·2–96·9), positive predictive value 95·7% (93·9–97·2), and negative predictive value 87·1% (84·2–89·6; table 2).

The external validation dataset comprised 246 people with suspected epilepsy who passed at least one screening criterion. 84 were identified as controls (no diagnosis of epilepsy), and 162 were cases (confirmed diagnosis of convulsive epilepsy). There were 45 females and 39 males in the control cohort (median age 51, IQR 16–85) and 80 females and 82 males in the case cohort (median age 27, 10–44). Of the cases of convulsive epilepsy, 57·4% had generalised seizures, 25·9% focal, and 16·7% combined focal and generalised seizures. The model was used to predict convulsive epilepsy in this cohort with the same probability threshold. The AUC was 0·95 (95% CI 0·92–0·98) whereas the sensitivity, specificity, and positive and negative predictive values were 97·5% (93·7–99·3), 82·4% (71·2–90·5), 92·9% (87·9–96·3), and 93·3% (83·8–98·2; table 2).

We implemented the logistic regression coefficients into the EDC. We adapted the weights from the logistic regression model into a paper-based questionnaire (The Oxford Convulsive Epilepsy Screen; appendix p 6). We tested the EDC in every Android API version from 16 to 30 on several devices and confirmed multidevice and version compatibility.

	AUC	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Internal validation	0·92 (0·91–0·94)	85·0 (81·8–87·7)	93·7 (91·5–95·4)	92·6 (90·1–94·7)	86·9 (84·1–89·3)
LOSO	0·94 (0·93–0·96)	88·2 (85·6–90·5)	95·3 (93·2–96·9)	95·7 (93·9–97·2)	87·1 (84·2–89·6)
External validation	0·95 (0·92–0·98)	97·5 (93·7–99·3)	82·4 (71·2–90·5)	92·9 (87·9–96·3)	93·3 (83·8–98·2)

Data in brackets are 95% CI. The internal-validation dataset contained a balanced sample of data from each clinical site. The LOSO model was trained using the data from all sites except Kilifi, Kenya, which was instead used for testing the model. The external-validation dataset comprised prospectively collected data from Kilifi, Kenya from people with suspected epilepsy who passed at least one of the two screening questions. AUC=receiver operator characteristic area under the curve. LOSO=leave one site out. PPV=positive predictive value. NPV=negative predictive value.

Table 2: **Evaluation of the logistic regression model for predicting convulsive epilepsy**

Discussion

We have established core phenotypic features enabling better individual triage by non-physician health-care workers and facilitating more appropriate onward referral to a neurologist for diagnostic confirmation of convulsive epilepsy. We have shown that eight binary questions can identify convulsive epilepsy with high sensitivity and specificity in sub-Saharan African populations. We confirmed this performance across five sites using internal validation data and subsequent external validation.

Of the eight selected features, one (tongue biting) identifies the convulsive phase, two capture impaired consciousness (becoming dazed and loss of contact with surroundings), and two identify the clonic phase (becoming stiff and urinary incontinence).[21–23] Two features (visual aura and abdominal pain) suggest focal onset.[24] One feature (is the seizure provoked?) was negatively associated with the diagnosis of epilepsy, probably identifying symptomatic seizures.[25] The exclusion of olfactory aura might be explained by the relative rarity with which olfactory auras occur and their association with other neurological conditions (eg, migraine). The omission of limb or face involvement might be associated with focal epilepsies or the reduced reliability of reporting this symptom.[26] We analysed the two screening questions from the first stage of the SEEDs study and confirmed they would have been removed during feature selection because of substantial multicollinearity. We instead used these questions as preselection criteria in the EDC.

We have developed a predictive model and free clinical app to screen people who present with potential symptoms of convulsive epilepsy. Entry of discriminating features, which we converted to simple clinical questions within this tool, requires minimal clinical training of community health-care workers. The EDC lends itself to rapid remote deployment and primary care physicians, nurses, pharmacists, and other health-care workers can learn to use the EDC with inbuilt supportive educational media.

The simple screening of individuals in these regions can lead to a substantial reduction in the diagnostic gap.[27] Identifying people who might have convulsive epilepsy has several benefits. First, this identification will reduce the risk to the individual by enabling appropriate safety education and mitigation of potential provoking factors of seizures. Second, by increasing the appropriateness of onward referral, the care pathway is streamlined, which can be cost saving at a societal level.[28] Although additional work is required, we hope that the app might help dispel some of the stigmatisation of epilepsy by providing estimated outputs confirming an organic diagnosis and the opportunity for further education.[29] Metadata from the EDC could also be used to update knowledge of disease prevalence in these regions without requiring resource-intense studies. We would also envisage

reducing the burden on health-care resources by empowering community health-care workers in sub-Saharan Africa, which is essential given the high prevalence of convulsive epilepsy in this region.

We propose that to implement predictive clinical tools in LMICs successfully, at least three key criteria must be met: robust, transparent modelling of large region-specific and culture-specific datasets; adequate validation using data from these areas; and permanent free and open access to these clinical technologies with iterative improvements informed by testing the predictive model and app in local communities.

Diagnostic app technology must be grounded in the population it seeks to serve. Although important as proof-of-concept work, previous studies have tended to use small, geographically limited datasets to generalise to a substantially larger and more diverse population. Those models have underexplored robust model development, do not have adequate validation, require more features than the EDC, under-report on diagnostic performance, and do not have discrete diagnostic classifications.[30,31] Some of those clinical tools have also introduced paywalls, potentially limiting access for people in LMICs. Although there will be similarities, deployment of apps developed elsewhere in the world to Africa might create unintended risks through questions being less specific and sensitive, being inappropriately phrased, or being incorrectly weighted.

Although the requirement for clinical training to administer such questionnaire panels is markedly lower than other clinical tools, adequate training remains paramount. Smartphone and tablet devices are increasingly ubiquitous in sub-Saharan Africa, alongside rapidly improving network infrastructure. Deployment of these clinical tools via smartphones or tablets can enable remote training of community health-care workers via the app itself.

The principal limitations of the EDC are that it is not applicable to non-convulsive seizures. We could not develop a diagnostic aid that accounted for these other conditions because of the limitations of the current dataset. Work is ongoing to create large databases of people with non-convulsive epilepsy to expand our predictive model. An additional limitation of this diagnostic aid is that it has been developed and validated within regions from which the data were collected and has not yet been validated in other areas whose populations did not contribute to model development. There is likely to be a need to culturally contextualise future versions before deployment in different settings. We aim to use data collected from these other countries to perform further external validation of the EDC and iteratively improve diagnostic accuracy. Importantly, the probability threshold for a prediction of convulsive epilepsy should ideally be adapted to the context of the target population while adjustments to the model such as intercept correction should be considered. Whether to refer the person with suspected epilepsy to specialist medical services initially versus commencing immediate treatment should be carefully evaluated.

The EDC provides a powerful tool to help diagnose convulsive epilepsy in resource-poor settings. The aim is not to replace health professionals, but rather aid in directing the individual to see the most appropriate clinician. Field studies are being initiated to further validate the EDC, whereby the model's predictions will be compared to diagnoses made by local neurologists at additional clinical sites. This will enable further refinement of the predictive model and the app as an iterative process.

The EDC is based on data from African sites. Although it can form the template for apps in other LMICs, similar groundwork should be done in different areas to contextualise the algorithm and app for the communities in which it is to be applied. Work is also needed to better understand non-convulsive seizures and psychogenic non-epileptic seizures such that similar diagnostics can be developed for these conditions in LMICs.

The EDC, a predictive model and app to identify convulsive seizures, has been rigorously developed and validated on a large dataset of African individuals. This bespoke tool might enable earlier and more accurate diagnosis leading to improved care and destigmatisation of people with convulsive epilepsy.

Contributors

GDJ contributed to the conception, methodology, statistical analysis, model development, validation, software (app) development, and wrote the original draft of the manuscript. SMK contributed to data curation, acquisition of external validation data, project administration, and reviewing and editing the manuscript. AKN, AKM, HM, SO-A, and RW contributed to data curation and reviewing and editing the manuscript. JHC, JWS, and CRN contributed to conception, funding acquisition, supervision, and reviewing the manuscript. CRN was Chief Investigator for the SEEDS study and contributed to the conception, funding acquisition, supervision, and reviewing of the manuscript. AS contributed to conception, funding acquisition, supervision, reviewing the manuscript, and the decision to submit for publication. No authors were precluded from accessing the data.

Declaration of interests

We declare no competing interests.

Data sharing

We welcome collaborations. Given the multisite origin of these data, requests for data will require approval from the clinical sites and partner institutions; requests can be made to the corresponding author. Queries regarding access to the data in the event the corresponding author is no longer available for contact should be addressed to Charles R Newton (charles.newton@psych.ox.ac.uk).

Acknowledgments

This work was supported by The Alan Turing Institute's Enrichment Scheme. JWS is based at the National Institute of Health Research (NIHR) University College London (UCL) Hospitals Biomedical Research Centre, which receives funding from the UK Department of Health's Research Centres funding scheme. JWS receives research support from the Dr Marvin Weil Epilepsy Research Fund and from the UK Epilepsy Society and the Christelijke Vereniging voor de Verpleging van Lijders aan Epilepsie, Netherlands. JWS or his department have also received grants from Esai, Angelina Pharma, and UCB Pharma. The UK Epilepsy Society endows the current position of JWS and he serves as Editorial Board member for *The Lancet Neurology*. JHC's research is supported by the NIHR Biomedical Research Centre at Great Ormond

Street Hospital. JHC holds an endowed chair at UCL Great Ormond Street Institute of Child Health, has received grants from the NIHR, Engineering and Physical Sciences Research Council, Great Ormond Street Hospital Children's Charity, Epilepsy Research UK, the Waterloo Foundation, and the Great Ormond Street Hospital Biomedical Research Centre. AS is supported by the Oxford Biomedical Research Centre. This work is funded by the NIHR and the Oxford Martin Programme on Global Epilepsy. This research was commissioned by the National Institute for Health Research using Official Development Assistance funding. The views expressed in this publications are those of the authors and not necessarily those of the UK National Health Service, the NIHR, nor the Department of Health and Social Care. The Wellcome Trust, UK, funded the SEEDS study through a Wellcome Trust Senior Fellowship in Clinical Tropical Medicine (083744) and the EPInA project is funded by the UK NIHR. This research was commissioned by the National Institute for Health Research (NIHR200134) using Official Development Assistance funding. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the department of Health and Social Care.

References

1 WHO. Epilepsy. 2019. https://www.who.int/news-room/fact-sheets/detail/epilepsy (accessed June 1, 2022).
2 Newton CR, Garcia HH. Epilepsy in poor regions of the world. *Lancet* 2012; **380:** 1193–201.
3 Thurman DJ, Begley CE, Carpio A, et al. The primary prevention of epilepsy: a report of the Prevention Task Force of the International League Against Epilepsy. *Epilepsia* 2018; **59:** 905–14.
4 Dorsey ER, Glidden AM, Holloway MR, Birbeck GL, Schwamm LH. Teleneurology and mobile technologies: the future of neurological care. *Nat Rev Neurol* 2018; **14:** 285–97.
5 Ngugi AK, Bottomley C, Kleinschmidt I, et al. Prevalence of active convulsive epilepsy in sub-Saharan Africa and associated risk factors: cross-sectional and case-control studies. *Lancet Neurol* 2013; **12:** 253–63.
6 Ngugi AK, Bottomley C, Chengo E, et al. The validation of a three-stage screening methodology for detecting active convulsive epilepsy in population-based studies in health and demographic surveillance systems. *Emerg Them Epidemiol* 2012; **9:** 1–8.
7 Placencia M, Sander J, Shorvon S, Ellison R, Cascante S. Validation of a screening questionnaire for the detection of epileptic seizures in epidemiological studies. *Brain* 1992; **115:** 783–94.
8 Rwiza H, Kilonzo G, Haule J, et al. Prevalence and incidence of epilepsy in Ulanga, a rural Tanzanian district: a community-based study. *Epilepsia* 1992; **33:** 1051–56.
9 Pal DK, Das T, Sengupta S. Comparison of key informant and survey methods for ascertainment of childhood epilepsy in West Bengal, India. *Int J Epidemiol* 1998; **27:** 672–76.
10 Edwards T, Scott AG, Munyoki G, et al. Active convulsive epilepsy in a rural district of Kenya: a study of prevalence and possible risk factors. *Lancet Neurol* 2008; **7:** 50–56.
11 Shannon CE. A mathematical theory of communication. *Bell Syst Tech J* 1948; **27:** 379–423.
12 Rosenfeld A, Graham DG, Jevons S, et al. Development and validation of a risk prediction model to diagnose Barrett's oesophagus (MARK-BE): a case-control machine learning approach. *Lancet Digit Health* 2020; **2:** e37–48.
13 Hall MA. Correlation-based feature selection for machine learning. PhD thesis, University of Waikato, 1999.
14 James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning. New York: Springer, 2013.
15 Moturu ST, Johnson WG, Liu H. Predicting future high-cost patients: a real-world risk modeling application. 2007 IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2007); Nov 2–4, 2007 (abstr 9882707).
16 Zhang W, Zeng F, Wu X, Zhang X, Jiang R. A comparative study of ensemble learning approaches in the classification of breast cancer metastasis. 2009 International Joint Conference on Bioinformatics, Systems Biology and Intelligent Computing; Aug 3–5, 2009 (abstr 10891490).
17 Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006; **26:** 565–74.
18 GlobalStats S. Mobile operating system market share Africa 2021. https://gs.statcounter.com/os-market-share/mobile/africa (accessed Sept 30, 2022).
19 Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD). The TRIPOD statement. *Circulation* 2015; **131:** 211–19.
20 Huang Y, Pepe MS. Assessing risk prediction models in case-control studies using semiparametric and nonparametric methods. *Stat Med* 2010; 29: 1391–410.
21 Devinsky O. Effects of seizures on autonomic and cardiovascular function. *Epilepsy Curr* 2004; **4:** 43–46.
22 Oliva M, Pattison C, Carino J, Roten A, Matkovic Z, O'Brien TJ. The diagnostic value of oral lacerations and incontinence during convulsive "seizures". *Epilepsia* 2008; **49:** 962–67.
23 Brigo F, Nardone R, Bongiovanni LG. Value of tongue biting in the differential diagnosis between epileptic seizures and syncope. *Seizure* 2012; **21:** 568–72.
24 Bien CG, Benninger FO, Urbach H, Schramm J, Kurthen M, Elger CE. Localizing value of epileptic visual auras. *Brain* 2000; **123:** 244–53.
25 Gunawardane N, Fields M. Acute symptomatic seizures and provoked seizures: to treat or not to treat? *Curr Treat Opt Neurol* 2018; **20:** 1–15.
26 Acharya V, Acharya J, Lüders H. Olfactory epileptic auras. *Neurology* 1998; **51:** 56–61.
27 Paddick S-M, Yoseph M, Gray WK, et al. Effectiveness of app-based cognitive screening for dementia by lay health workers in low resource settings. A validation and feasibility study in rural Tanzania. *J Geriatr Psychiatr Neurol* 2020; **1:** 0891988720957105.
28 Iribarren SJ, Cato K, Falzon L, Stone PW. What is the economic evidence for mHealth? A systematic review of economic evaluations of mHealth solutions. *PLoS One* 2017; **12:** e0170581.
29 Ekstrand ML, Raj T, Heylen E, et al. Reducing HIV stigma among healthcare providers in India using a partly tablet-administered intervention: the DriSti trial. *AIDS Care* 2020; **32** (suppl 2)**:** 14–22.
30 Giuliano L, Cicero CE, Trimarchi G, et al. Usefulness of a smartphone application for the diagnosis of epilepsy: validation study in high-income and rural low-income countries. *Epilepsy Behav* 2021; **115:** 107680.
31 Patterson V, Pant P, Gautam N, Bhandari A. A Bayesian tool for epilepsy diagnosis in the resource-poor world: development and early validation. *Seizure* 2014; **23:** 567–69.