Edinburgh Research Explorer

Case-control study developing Scottish Epilepsy Deaths Study score to predict epilepsy-related death

Published In:
Brain

Case-control study developing Scottish Epilepsy Deaths Study score to predict epilepsy-related death

Gashirai K. Mbizvo,[1,2] Christian Schnier,[3] Colin R. Simpson,[3,4] Susan E. Duncan[1,5,†] and Richard F. M. Chin[1,6,†]

†These authors contributed equally to this work.

Abstract

This study aims to develop a risk prediction model for epilepsy-related death in adults.

In this age- and sex-matched case-control study, we compared adults (aged ≥16 years) who had epilepsy-related death between 2009–2016 to living adults with epilepsy in Scotland. Cases were identified from validated administrative national datasets linked to mortality records. ICD-10 cause-of-death coding was used to define epilepsy-related death. Controls were recruited from a research database and epilepsy clinics. Clinical data from medical records were abstracted and used to undertake univariable and multivariable conditional logistic regression to develop a risk prediction model consisting of four variables chosen *a priori*. A weighted sum of the factors present was taken to create a risk index – the Scottish Epilepsy Deaths Study Score (SEDS Score). Odds ratios (OR) were estimated with 95% confidence intervals (CIs).

224 deceased cases (mean age 48 years, 114 male) and 224 matched living controls were compared. In univariable analysis, predictors of epilepsy-related death were recent epilepsy-related accident and emergency (A&E) attendance (OR 5.1, 95% CI 3.2–8.3), living in deprived areas (OR 2.5, 95% CI 1.6–4.0), developmental epilepsy (OR 3.1, 95% CI 1.7–5.7), raised Charlson Comorbidity Index (CCI) score (OR 2.5, 95% CI 1.2–5.2), alcohol abuse (OR 4.4, 95% CI 2.2–9.2), absent recent neurology review (OR 3.8, 95% CI 2.4–6.1), and generalised epilepsy (OR 1.9, 95% CI 1.2–3.0). SEDS Score model variables were derived from the first four listed above, with CCI ≥2 given 1 point, living in the two most deprived areas given 2 points, having an inherited or congenital aetiology or risk factor for developing epilepsy given 2 points, and recent epilepsy-related A&E attendance given 3 points. Compared to having a SEDS Score of 0, those with a SEDS Score of 1 remained low risk, with OR 1.6 (95% CI 0.5–4.8). Those with a

SEDS Score of 2–3 had moderate risk, with OR 2.8 (95% CI 1.3–6.2). Those with a SEDS Score of 4–5 and 6–8 were high risk, with OR 14.4 (95% CI 5.9–35.2) and 24.0 (95% CI 8.1–71.2), respectively.

The SEDS Score may be a helpful tool for identifying adults at high risk of epilepsy-related death and requires external validation.

Author affiliations:

1 Muir Maxwell Epilepsy Centre, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK

2 Pharmacology and Therapeutics, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, UK

3 Usher Institute, University of Edinburgh, Edinburgh, UK

4 School of Health, Wellington Faculty of Health, Victoria University of Wellington, Wellington, New Zealand

5 Department of Clinical Neurosciences, NHS Lothian, Edinburgh, UK

6 Royal Hospital for Children and Young People, Edinburgh, UK

Correspondence to: Gashirai Mbizvo

Muir Maxwell Epilepsy Centre, University of Edinburgh, 50 Little France Crescent Edinburgh BioQuarter, EH16 4TJ, UK

E-mail: Gashirai.Mbizvo@ed.ac.uk

Running title: The SEDS Score

Keywords: cause of death; seizures; risk prediction modelling; routine data; terminal illness

Abbreviations: A&E = Accident and emergency; ASMs = Antiseizure medications; CCI – Charlson Comorbidity Index; CI = Confidence interval; G40 = ICD-10 code for epilepsy; G41 = ICD-10 code for status epilepticus; GTCS = Generalised tonic-clonic seizure; PBPP = Public Benefit and Privacy Panel for Health and Social Care; ICD = International Classification of

Diseases; NHS = National Health Service; NPV = Negative predictive value; OR = Odds ratio; PPV = Positive predictive value; PWE = People with epilepsy; R56.8 = ICD-10 code for seizures; RPM = Risk prediction model; SAIL = Secure Anonymised Information Linkage; SEDS Score = Scottish Epilepsy Deaths Study Score; SIMD = Scottish Index of Multiple Deprivation; SHARE = Scottish Health Research Register; SUDEP = Sudden unexpected death in epilepsy; TIA = Transient ischaemic attack

1.1 Introduction

Epilepsy is common, affecting 70 million people globally.[1] In the United Kingdom, there are 600,000 people with epilepsy (PWE), of which 60,000 live in Scotland.[2] PWE are at significantly increased risk of premature death.[3] Some of those deaths may be entirely unrelated to their epilepsy. However, a substantial proportion relate to the epilepsy itself, its treatment, or comorbidities associated with epilepsy.[3] Together, these are termed 'epilepsy-related deaths', and they are likely to contribute a significant burden of total years of potential life lost.[3,4] A recent study of 1,921 epilepsy-related deaths identified across Scotland between 2009–2016 showed that these have remained high over time, with age-specific mortality ratios significantly increased in young adults aged 16–54 years, peaking at 5.3 (95% confidence interval (CI) 1.8–8.8) in those aged 16–24 years.[5] Sudden unexpected death in epilepsy (SUDEP) constituted 30% of the 553 young adult epilepsy-related deaths, with several other non-SUDEP fatal mechanisms identified including aspiration pneumonia, cardiac arrest, antiseizure medication (ASM) or narcotic poisoning, drowning, and alcohol dependence. A substantial number of these young adult epilepsy-related deaths were potentially avoidable,[5] as has also been suggested in other countries.[6-8] This highlights the importance of trying to identify which individuals are at highest risk so they can be targeted with earlier and more aggressive care. Few studies have attempted to do this through risk prediction modelling (RPM), particularly focusing on epilepsy-related deaths specifically.[9-14] To help address this, we have undertaken a Scottish multi-centre case-control study using routine clinical information gathered from medical records. The study aims to develop a pragmatic RPM to identify adults at increased risk of epilepsy-related death. We will describe multiple potential risk factors for epilepsy-related death in order to help build the model.

3

1.2 Materials and methods

1.2.1 Study design summary

We carried out an age- and sex-matched case-control study comparing adults who had epilepsy-related death to living adults with epilepsy. Data pertaining to potential risk factors for epilepsy-related death were collected from the medical records of cases and controls to allow multivariable analysis and development of a RPM. The study was designed for model development and internal validation,[15] and follows standard reporting guidelines.[16]

1.2.2 Data sources, and participants

1.2.2.1 Case ascertainment

Cases were screened for inclusion from the dataset of a prior study investigating the diagnostic accuracy of four administrative healthcare datasets used to identify deceased adults with epilepsy in Scotland.[17] The study linked national mortality records, hospital admissions, ASM prescriptions, and regional primary care attendances to identify adults (aged ≥16 years) who died between 01/01/09–01/01/16. Coded indicators of epilepsy were compared to confirmed epilepsy diagnoses made by a senior epileptologist upon reviewing medical records (*see* Confirming epilepsy diagnosis section). The study identified 614 confirmed epilepsy cases who had epilepsy-related death during the 2009–16 study period. Epilepsy-related death was defined as death occurring with at least one ICD-10-coded cause listed as G40–41 (epilepsy–status epilepticus) and/or R56.8 (seizure) in any position within the sequence of events that lead to death in the mortality records, or with cause of death free texts pertaining to seizures (alongside at least one community-prescribed ASM), epilepsy, or SUDEP.[5,17] Such a definition is consistent with national guidance that the presence of an epilepsy code/indicator anywhere in the cause of death record indicates that epilepsy was thought by the certifying doctor to be either part of the sequence of events leading to death or else a factor contributing to the death.[18,19] Additional case-ascertainment benefit was drawn from the routine cause-of-death quality assurance processes in place to improve the accuracy of Scottish death certification and coding,[20,21] the robustness of which was highlighted recently.[22] These quality assurance processes helped reduce the likelihood of there being an underestimation in the number of epilepsy-related deaths captured by the study,[5,17] which can sometimes happen in studies using death certification.

4

1.2.2.2 Control ascertainment

Living control participants were screened for inclusion from a nationwide Scottish Health Research Register (SHARE),[23,24] supplemented by controls recruited directly from NHS Lothian epilepsy clinics. Written consent was obtained from both groups.

SHARE controls: The SHARE register consists of nearly 300,000 research volunteers in Scotland, aged ≥11 years.[23,24] The SHARE research staff connected us with SHARE registrants aged ≥16 years with at least one G40–41-ICD-10-coded activity in their health records and consenting to be contacted by us. We reviewed the medical records of those consenting to participate.

Clinic controls: Clinic controls were recruited in a consecutive manner directly from NHS Lothian epilepsy clinics. We reviewed the medical records of those consenting to participate.

Eligible cases and controls required sufficiently complete medical records to allow confirmation of a diagnosis of epilepsy and abstraction of clinical data. Controls were matched 1:1 by age (± 10 years) and sex to cases of epilepsy-related death after recruitment.

1.2.2.3 Confirming epilepsy diagnosis

Access to medical records was made possible through the support of a national network of clinical colleagues, as described in greater detail previously.[5,17] For each candidate case and control, the medical records were reviewed by an experienced consultant epileptologist (S.E.D.), who used the medical records to confirm the presence or absence of epilepsy.[25] This was based on corroborative evidence such as the presence of two or more unprovoked seizures or clear documentation of a diagnosis of epilepsy from a neurologist.[26,27] The medical records were extensive, and included electronic and paper records from general and/or specialist inpatients, emergency care, outpatients (including neurology clinics), hospital discharge summaries, referral letters from general practitioners, radiology scans, neurophysiology reports, and medication lists.

1.2.2.4 Data abstraction, outcome and predictors

The outcome for model prediction fell into binary categories: 1 = deceased (i.e. epilepsy-related death) and 0 = alive (living control).

Abstraction of data from medical records was undertaken by S.E.D. during several hospital visits made between 11/10/2017–24/02/2019 for cases, and 25/02/2019–15/01/2020 for controls. Data

5

on multiple potential risk factors were collected using a data abstraction tool developed *a priori* using systematic review evidence (Appendix S1).[3] Candidate predictor variable selection was refined by consensus amongst the research team and informed by consultation with external epilepsy specialists, patient and charity groups, and parliamentary policymakers.[28-30] Table 1 lists the predictor variables that were ultimately selected as candidates for model development. Scottish Index of Multiple Deprivation (SIMD) looks at the extent to which an area is deprived across seven domains: income, employment, education, health, access to services, crime and housing. This is ranked from 1 = most deprived to 5 = least deprived areas of residence.[31] For each control participant, we inserted their home postcode into the Scottish Government's publicly-available SIMD postcode lookup file to reveal their SIMD quintile.[31] Charlson Comorbidity Index (CCI) is a well-validated prognostic index of comorbid conditions that alter 10-year survival in patients with multiple comorbidities.[32-34] The comorbidities captured include cardiovascular, cerebrovascular, respiratory, neurodegenerative, and neoplastic. CCI is scored from 0–37 points, with higher scores associated with lower survival. We extracted each participant's current comorbidities from the medical records and used this to calculate their CCI score, which is easily done using an online calculator (www.mdcalc.com/calc/3917/charlson-comorbidity-index-cci).[35] We grouped the CCI scores into categories 0–1 (associated with a 96–98% estimated 10-year survival), 2–3 (77–90% estimated 10-year survival), and ≥4 (≤53% estimated 10-year survival),[32-34] as grouped elsewhere.[36] CCI scores were used as they would allow us to capture the comorbidities most relevant to mortality and increase applicability of our methods in other regions, as CCI is widely used.[36-38]

1.2.3   Statistical analysis

All analyses were performed using IBM SPSS Statistics for Macintosh, Version 25.0. Armonk, NY: IBM Corp. Where relevant, categories with less than five events/participants were combined or reported as '<5' to avoid patient identification.[39]

1.2.3.1 Sample size

We estimated that we could capture a convenience sample size of 200–250 controls, matched to a mirrored number of age- and sex-matched cases taken from the 614 deceased adults who had epilepsy-related death, for whom medical records had already been reviewed during the diagnostic accuracy study.[17] This number of controls was based on our estimated capacity and

the time taken to fully recruit living controls and abstract their clinical data from medical records. Using a recommended rule of 50 events per variable for logistic regression to estimate coefficients with adequate precision,[40,41] such a sample size would allow for a multivariable RPM size of four predictor variables. Using more variables than the sample data could support would increase the risk of overfitting (achieving overly optimistic predictive accuracy and hence resulting in failure to replicate the results elsewhere).[42]

1.2.3.2 Missing data

An *a priori* threshold was set to exclude from further analysis any variables with ≥30% of their values missing either within cases or within controls, based on prior literature.[43] We also excluded any variables with >10% difference in the missing data burden between cases compared to controls.[44] These exclusions allowed us to prioritise creating a pragmatic RPM consisting of variables likely to be found complete in medical records, and that were uniformly represented between cases and controls. For variables with <30% of their values missing, an overall summary of the missing data patterns was given and a logistic regression model for multiple imputation was used to replace the missing values. The missing values were imputed using fully conditional specification techniques based on the iterative Markov chain Monte Carlo (MCMC) method, with 20 imputed datasets gathered and imputed results pooled.[45,46] The missing data handling methods used were based on an assumption data were missing at random (MAR). This was a plausible assumption given that a large number of clinically relevant variables were tested, meaning the probability of missingness would likely depend on only the observed data already captured within the dataset.[47]

1.2.3.3 Model development

Where possible, we reported both the original and imputed datasets for the models developed. Primary analysis was of the original data (complete case analysis), and the imputed data were used for sensitivity analysis.

Univariable analysis: Univariable analysis was performed using conditional logistic regression for included variables to provide an odds ratio (OR) and 95% CI.

Correlation: Correlation was examined in a pairwise fashion between the variables, using contingency tables and Chi-square tests with a 2-sided 5% significance. Where correlation was

significant (p < 0.05 for both complete case and imputed data), the Lambda statistic was used to assess the strength of association, interpreted as 0.01–0.09 = weak, 0.10–0.29 = moderate, 0.30–0.99 = strong, and 1.00 = perfect association.[48] Correlation was assessed for both cases and controls separately, with the highest Lambda statistic between both groups reported (Table 2).

Model selection: Two consultant epileptologists (S.E.D. and R.F.M.C.) independently made a clinically-driven decision to select four variables for inclusion in the final RPM,[42] whilst blinded to the univariable results and taking into account variable correlations and missing data burdens (see Appendix table S2 for justification of each choice made). Any disagreements in variable selection were resolved by consensus. This variable selection approach was undertaken because it is often advised that variable selection should be focused more on clinical knowledge and previous literature than statistical selection methods alone.[42,49] The rationale is that it is better to select variables based on a wider body of clinical knowledge than to try to depend on statistical significance of results from a lesser quantity of information, particularly in a moderate-sized sample such as ours.[50,51] A data-driven, sample-based, selection would be more sensitive to random variation in the data points due to sampling variability and more likely to generate false signals. Furthermore, there are several problems with the popular statistical selection methods themselves, including stepwise variable selection, backward elimination, and forward selection, as summarised in detail elsewhere.[50]

Multivariable analysis: Multivariable conditional logistic regression was performed to provide ORs for the four variables selected for risk prediction modelling, adjusted for one another, alongside 95% CIs. Any variables (or levels within variables) retaining significance within the univariable analysis were taken forward into the multivariable model, where they were compared against all remaining cases as a reference for nominal variables. For ordinal variables, any lower-ranking levels that were positive were combined with higher-ranking levels to create a positivity threshold. We assessed the positive variables within the multivariable model and assigned a weight (number of points) to each, based on the observed magnitudes of ORs in the complete case analysis, but also taking imputed data trends into account as well.[52] A sum of these points was then taken to give a total score – which we termed a Scottish Epilepsy Deaths Study score (SEDS Score). An internal validation process was performed for each model (i.e. quantifying statistical performance of the model using the data on which the model was developed).[15] This was done by performing 1,000 bootstrap analysis samples to provide more robust p-values,

1 which were then checked for agreement against the original model's p-values for each variable.

2 Where there was poor agreement, the bootstrap result was prioritised.

3 We used Nagelkerke's R^2 to estimate what proportion of variance in the outcome was explained

4 by the predictor variables in the model, and the Hosmer-Lemeshow goodness-of-fit test to

5 determine whether the model adequately described these outcome data (where $p \geq 0.05$ indicates

6 a good fit). The Hosmer-Lemeshow contingency table was used to assess model calibration by

7 illustrating agreement between subgroups of predicted probabilities and their observed

8 frequencies.[15,53] Discrimination describes how well the model differentiates between patients

9 who experienced the outcome (deceased cases) and those who did not (living controls).[15] This

10 would normally be measured using an area under the receiver operating characteristic (ROC)

11 curve.[53] However, ROC statistics can be misleading for binary or categorical predictors and, as

12 such, they are cautioned against in this scenario.[54,55] Therefore, we approximated discrimination

13 primarily through indicating the model's ability to capture all of the deceased cases by estimating

14 its sensitivity, and all of the living outcomes by estimating its specificity. Positive predictive

15 value (PPV), negative predictive value (NPV), and model accuracy (sensitivity $\times$ prevalence +

16 specificity $\times$ (1 − prevalence)) were reported for reference only given that these estimates depend

17 on prevalence of the outcome, which was fixed at 50% in this study because of the 1:1 matched

18 case-control design. 95% CIs were included.

19 1.2.4   Approvals

20 This study was approved by South East Scotland Research Ethics Committee 2 (15/SS/0165,

21 IRAS 181131), and the Scottish Public Benefit and Privacy Panel for Health and Social Care

22 (PBPP).

Data availability

24 Approved researchers can access linked datasets via application to the electronic Data Research

25 and Innovation Service (eDRIS) of the Scottish Information Services Division (ISD) at

26 www.isdscotland.org/Products-and-Services/eDRIS//. Aggregate case-control data is available

27 upon reasonable request (www.muirmaxwellcentre.com/contact-us/). For the purpose of open

28 access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author

29 Accepted Manuscript version arising from this submission.

1

1.3 Results

1.3.1 Baseline characteristics of cases and controls

The study cohort comprised of 224 deceased cases (114 male, 110 female) and 224 living controls (114 males, 110 female), spread across multiple NHS Health Board regions in Scotland (see Figure 1). Mean age at death for cases was 48 years (±17 SD, range 16–90). This was no different to the mean age at assessment of 48 years (± 16 SD, range 19–87) for controls, demonstrating that the age-matching limits of ± 10 years did not result in any major differences in age between the cases and controls that were actually studied. 124 controls were recruited from SHARE Scotland,[23] with the remaining 100 recruited from NHS Lothian epilepsy clinics.

1.3.2 Missing data

Several predictor variables were excluded from further analysis owing to their missing data burden (see Table 1 and Appendix table S2). The remaining variables were taken forward into multiple imputation and univariable analysis.

1.3.3 Correlation

A strong association was seen only between the type of seizures (focal or general) and the risk factors for developing epilepsy (inherited/congenital or acquired) – see Table 2. The remaining pairs demonstrated moderate or weak association, with the greatest association amongst these being between having been seen by a neurologist in the preceding year and the number of ASMs prescribed when last seen.

1.3.4 Univariable analysis

The odds of epilepsy-related death occurring were significantly increased in those who lived in the two most deprived SIMD quintiles, those with a history of alcohol abuse in the preceding two years, those with generalised epilepsy, those who had experienced at least one accident and emergency (A&E) attendance or hospital admission in the preceding year related to seizures or epilepsy, those with evidence of mental health problems in the preceding two years, those who had not been seen by a neurologist in the preceding year, those with a CCI of 2–3, and those with an inherited or congenital aetiology for developing epilepsy (see Table 3). ORs and CIs for CCI

10

1 2–3 and ≥4 almost entirely overlapped, indicating a similar level of association beyond CCI ≥2

2 in this dataset. The odds of epilepsy-related death occurring were significantly reduced in those

3 who had undergone previous epilepsy surgery in the complete case analysis. However, only a

4 small number of cases and controls had had epilepsy surgery (<10% of the sample, overall), and

5 the difference associated with this variable was lost when imputed data were included. The

6 number of ASMs taken when last seen were not significant individual predictors. Very few cases

7 and controls (<3% of the sample) were not on an ASM.

8 In an exploratory post-hoc analysis, the odds of epilepsy-related death occurring remained

9 significantly increased in those who had not been seen by a neurologist in the preceding year

10 even with analysis restricted to using only the SHARE control participants. This indicates that

11 this variable was significant even when excluding the sample of control participants recruited

12 directly from epilepsy clinics.

13 1.3.5   Multivariable analysis

14 The four variables ultimately selected for multivariable analysis are listed below (see Table 1 for

15 expanded definitions of each variable):

16 • SIMD Quintile;

17 • A&E attendance or hospital admission for epilepsy/seizures;

18 • Aetiology and risk factors for developing epilepsy;

19 • CCI.

20 Table 4 summarises results of the multivariable model derived from these four variables. The

21 SEDS Score will be calculated by taking a sum of these factors, when present, weighted as:[52]

22 • *1 point for a CCI of ≥2:* A CCI score ≥2 was given the lowest weighting of 1 as it had an OR

23 of only 1.6 in complete case analysis (which was not significant), but an OR of 2.6 in

24 multiple imputation analysis (which was significant). Our clinical interpretation of this,

25 supported by prior literature,[32-34] would suggest higher CCI scores would normally be

26 associated with increased mortality; suggesting some of the missing data within our CCI

27 variable could have contributed to the lower complete case result. However, we applied a

cautious weighting of only 1 as we did not wish to inflate the weight for this variable given the complete case result, which was the basis for primary analysis.

- *2 points for having inherited or congenital aetiology or risk factors for developing epilepsy:* assigned this weight because the complete case OR for this variable was higher than for CCI ≥2, was in a similar range to OR for the two most deprived SIMD quintiles, and was lower than OR for A&E attendance or hospital admission.

- *2 points for living in the two most deprived SIMD quintiles:* assigned this weight because complete case OR for this variable was similar to OR for inherited or congenital aetiology for developing epilepsy;

- *3 points for at least one seizure- or epilepsy-related A&E attendance and/or hospital admission in the preceding year:* assigned this weight because this variable had the highest complete case OR.

Therefore, each individual could have a SEDS Score of 0–8.

Model diagnostics showed evidence of internal validity for this multivariable model, with the *P* values from 1,000 bootstrap samples matching the significance patterns of the complete case analysis. The model was a good fit for the data (*P* > 0.05 for the Hosmer-Lemeshow goodness-of-fit test both using complete case and imputed data). 30% of the variance in the outcome was explained by the predictor variables included in this model (R^2 = 0.303). The model demonstrated good calibration, with the predicted estimates mirroring the actual estimates closely throughout the Hosmer-Lemeshow contingency table (Table 5). In terms of discrimination, the model demonstrated a sensitivity of 72% (95% CI 65–78%) and specificity of 72% (95% CI 65–79%) in the complete case analysis, indicating that the majority of cases and controls were identified. There was little difference in these discrimination figures when imputed values were included. PPV, NPV, and model accuracy were 74% (95% CI 69–79%), 70% (95% CI 64–75%), and 72% (95% CI 67–77%), respectively.

The SEDS Scores were analysed in a multivariable model (SEDS Score model), using a score of 0 points as the reference category. This was compared to SEDS Scores in groups of 1, 2–3, 4–5, and 6–8. The results of this model are summarised in Table 6. The odds of epilepsy-related death occurring were not significantly increased in those with a SEDS Score of 1. The odds were then

1 significantly increased in the remaining SEDS Score groups. The suggestion of a dose-response

2 in increasing odds was seen, with OR increasing from 2.8, to 14.4, then 24.0 for SEDS Scores of

3 2–3, 4–5, and 6–8, respectively. However, there were also increasingly fewer cases and controls

4 as SEDS Score increased, which acted to widen the CIs and thereby reduce confidence about the

5 precise magnitude of odds as they increased. There was little difference between the complete

6 case and imputed data analysis for SEDS Score modelling.

7 Model diagnostics showed evidence of internal model validity for the SEDS Score, with the *P*

8 values from 1,000 bootstrap samples matching the significance patterns of the complete case

9 analysis. The model was also a good fit for the data (p > 0.05 for the Hosmer-Lemeshow

10 goodness-of-fit test both using complete case and imputed data). 28% of the variance in the

11 outcome was explained by the predictor variables included in this model (R^2 = 0.283). The

12 model also demonstrated good calibration, with the predicted estimates mirroring the actual

13 estimates closely throughout the Hosmer-Lemeshow contingency table (Table 5). In terms of

14 discrimination, the model demonstrated a sensitivity of 58% (95% CI 51–64%) and specificity of

15 84% (95% CI 78–88%) using complete case data. There was little difference in these

16 discrimination figures when imputed values were included. PPV, NPV, and model accuracy were

17 78% (95% CI 72–83%), 66% (95% CI 63–70%), and 71% (95% 66–75%), respectively.

1.4 Discussion

1.4.1 Summary of findings

20 In the first reported case-control study investigating epilepsy-related death in Scotland, data from

21 the medical records of 224 deceased cases and living controls from multiple regional centres

22 were analysed to identify several potential risk factors for epilepsy-related death and to develop a

23 pragmatic RPM for identifying those at the highest risk.

24 In terms of risk factors, univariable analysis revealed that eight variables were significantly

25 associated with increased risk of epilepsy-related death. These were having a recent A&E

26 attendance or hospital admission for seizures or epilepsy, a recent history of alcohol abuse,

27 absent recent neurology review, the presence of inherited or congenital aetiology for developing

28 epilepsy, living in the two most deprived Scottish quintile areas, generalised epilepsy, a recent

29 history of mental health problems, and having a raised comorbidity burden (CCI ≥2). ORs

13

1 ranged between 1.6–5.1 between these variables. Multivariable modelling was undertaken for the

2 positive factors within variables CCI (≥2 assigned 1 point), SIMD quintile (two most deprived

3 assigned 2 points); aetiology or risk factors for developing epilepsy (inherited/congenital

4 assigned 2 points), and A&E attendance or hospital admission for epilepsy/seizures (assigned 3

5 points). Modelling of the composite SEDS Scores derived from these points revealed three levels

6 of risk for epilepsy-related death occurring: low risk in those with a SEDS Score of 1 (ORs 1.6–

7 1.7), moderate risk for a SEDS Score of 2–3 (ORs 2.8–3.3), and high risk for a SEDS Score of

8 ≥4 (ORs 14.4–14.7 for SEDS Scores 4–5 and 24.0–29.5 for SEDS Scores 6–8). Wide CIs were

9 noted surrounding the high-risk groups, owing to increasingly fewer events as factors were

10 combined. This serves to justify proceeding to a larger study in a different country to externally

11 validate the model and narrow CIs further in the process.[56] Appendix S3 illustrates the potential

12 avenue to clinical implementation of the SEDS Score as an example risk scoring card for

13 frontline clinicians to use.

14 1.4.2   Interpretation

15 Aside from a history of inherited or congenital aetiology and risk factors for developing epilepsy,

16 the remaining three predictors used to generate the SEDS Score were potentially modifiable and

17 serve as a target for future risk prevention strategies clinically or from a wider epidemiological

18 perspective. Seizures are the most common neurological cause of unscheduled hospital

19 admission in the UK.[57] Whilst such admissions are a sign of poorer epilepsy control, they are

20 often clinically unnecessary and typically lead to little benefit for epilepsy management.[58,59] As

21 such, there are several studies proposing methods to reduce seizure-related admissions in

22 PWE.[59-61] This may be achieved through, for example, intensifying scheduled specialist and

23 community resources for PWE, or developing alternative care pathways.[61,62] Our study findings

24 support this need as we show that seizure-related hospital admissions were the strongest

25 predictor for subsequent epilepsy-related death. SIMD is modifiable from a population

26 perspective, including from government incentives targeted toward high-risk groups.[28]

27 Comorbidities have be managed using approaches such as multidisciplinary team clinics to help

28 reduce mortality.[63]

29 Whilst it is well established in literature that alcohol abuse, congenital or developmental

30 epilepsies, social deprivation, generalised epilepsy, and mental health problems are risk factors

14

for all-cause mortality in PWE,[3,7,64] we show that they can be further streamlined into predicting *epilepsy-related* mortality specifically. Furthermore, we test the role of recent A&E attendance or hospital admission for seizures and epilepsy on predicting epilepsy-related mortality, and the role of absent neurology review within a year on this outcome. There is little study of these two variables in epilepsy mortality literature.[3] Our data suggest the odds of epilepsy-related death occurring may be up to five-times increased in those attending A&E or admitted to hospital for seizures/epilepsy within a year, and up to three times increased in those not reviewed by a neurologist within a year. The SUDEP and Seizure Safety Checklist is a free evidence-based tool supporting clinicians in discussing risk with PWE, with over 1,000 clinician subscribers in the UK currently.[11] It lists risk factors linked to epilepsy mortality including (but not restricted to) SUDEP. However, it was not developed using regression modelling of the proposed risk factors against one another, but rather as expert consensus on risk factors identified in literature.[11] These risk factors include, for example, active seizures, generalised tonic-clonic seizures, status epilepticus, nocturnal seizures, poor medicine adherence, alcohol or substance misuse, and depression.[11] Following on from the findings of our study, it may be reasonable to consider including A&E attendance or hospital admission within the preceding year for seizures/epilepsy, or including absent neurology review within a year in future iterations of the SUDEP and Seizure Safety Checklist.

RPMs should be both internally and externally validated before being adopted into clinical practice.[15,56] The preferred approach for internal validation is bootstrapping, which we have undertaken.[15,56] The preferred approach for external validation uses independent data from a different location than the development data to re-assesses model performance.[56] This is the logical next phase of research for the SEDS Score, and could be undertaken using the national Secure Anonymised Information Linkage (SAIL) databank in Wales, for example.[65,66] Prognostic models can also be assessed for their impact in a clinical setting.[56] For example, mortality outcomes could be compared between patients managed by clinicians with and without access to the SEDS score for triaging.[56] Such an approach was taken in the STarTBack trial for back pain and the results showed a significantly larger reduction in disability as well as cost savings in the group receiving prediction-model-based care compared with controls.[56]

Our study develops a RPM using standard observational methods for this purpose.[15,56] It does not test interventions to alter the SEDS Score variables or overall score, nor the effect this would

15

1 have on mortality. This would require a clinical trial. Therefore, we are unable to make
2 recommendations to clinicians about how to manage their patients effectively once a high SEDS
3 Score is identified. Such management would need to be individualised, based on patient
4 circumstances, local policy, and available treatments. However, by identifying the high-risk
5 individuals, our RPM would support clinicians in triaging which patients need to be prioritised.
6 For example, in a non-specialist setting, where such patients often first present,[57,58] a high SEDS
7 Score could prompt re-discussion about seizure safety, signposting online resources for support
8 (such as www.epilepsy.org.uk/info), and rapid-access epilepsy clinic referral (rather than
9 reserving such clinics for first-fit patients alone). The potential impact of such approaches is well
10 recognised. For example, the asthma deaths review demonstrated that many deaths were likely to
11 be preventable simply through better proactive clinical management, such as arranging for
12 patients to be seen promptly after a hospital admission, and through following clinical
13 guidelines.[67] In a specialist setting, a high SEDS Score could prompt referral for epilepsy surgery
14 earlier than would have otherwise been considered, discussion in a multidisciplinary team
15 meeting, further medication reviews, or implementation of seizure alarms and nocturnal
16 supervision.

17 Our previous study corroborated a heavy burden of seizure-related A&E attendances and hospital
18 admissions in the period leading up to epilepsy-related death, with a substantial number of these
19 patients lacking any recent neurology review or follow-up.[5] Such A&E attendances and hospital
20 admissions therefore represent missed opportunities for creating effective pathways to specialist
21 care, a problem also highlighted in England.[58] As such, it would be important for the SEDS
22 Score to be available to non-specialist doctors including GPs, general physicians and emergency
23 care doctors, and for it to be designed in such a way that these non-specialists could use the tool.
24 This includes basing it on variables likely to have little missing data in these non-specialist
25 medical records. Support from local and national charities and organisations would help with
26 disseminating the model to non-specialist areas, as is being done by SUDEP Action for the
27 SUDEP and Seizure Safety Checklist.[11] We would also apply to have the SEDS Score included
28 in free online clinical scoring tool databases, such as MDCalc (www.mdcalc.com), meaning it
29 would be easily accessible globally.[35]

30 The methods used in our study can be translated to any country with unique patient identifiers
31 and linked administrative data or electronic health records (EHRs). For example, all patients

16

using NHS services in England benefit from a unique patient identifier (NHS number). A study was recently published evaluating a nationwide cohort of more than 54 million people in England using NHS-number-linked EHRs and administrative data.[68] Studies from outside of the UK have used a similar approach, including in the US,[69] Canada,[70] Sweden,[71] and South Africa.[72] Our methodology is focused on developing a generalisable RPM, selecting variables which are clinically and statistically appropriate to include, with little overall missing data. Consequently, the model identifies four risk factors which should be readily available within patient medical records and could even be obtained from administrative data, facilitating clinical implementation or research in other regions in future. Each variable has been successfully studied in other global regions using EHRs or administrative data.[3,57,73-76]

1.4.3 Study limitations

Limited sample size allowed development of only a four-variable RPM. This is, however, a similar model size to the three-item SUDEP-3 inventory – a RPM for SUDEP that was developed from a 7-item inventory (SUDEP-7).[9,10,12] Whilst a smaller sample allowed us to investigate each medical record in greater detail, it meant we were unable match cases and controls by centre, instead pooling them all into one Scottish centre. However, both the cases and controls were mainly recruited from Scotland's Central Belt region (Glasgow, Ayrshire, Falkirk, Edinburgh, Lothian and Fife, see Figure 1).

Controls were drawn from those signing up to SHARE and those attending epilepsy clinics, yet no such selection process applied to cases. This may have created a selection bias, especially towards controls being generally healthier than cases. This is unlikely to have influenced the conclusions we drew from SEDS Scoring as the OR figures for this were markedly high, and would therefore be unlikely to have been entirely attributable to differences in baseline characteristics of the groups. Furthermore, SHARE provided evidence their study population is representative of the general Scottish population in terms SIMD and ethnicity (see Appendix table S4) with percentage difference between the two being no more than 2% in each SIMD decile. 4% and 7% of the general population and SHARE population, respectively, are from ethnic minority backgrounds.

Although our study attempted to extract data on several other potentially modifiable risk factors for epilepsy-related death including seizure frequency, timing, seizure alarm use, and a history of

17

status epilepticus, large amounts of missing data in the medical records precluded analysis of these. We were strict to exclude from further analysis any variables with ≥30% missing data because including variables with large amounts of missing data would not only substantially reduce the model's predictive power,[43] but also it would curtail its clinical utility. This is because end users might also struggle to find these variables within their medical records. The resultant predictive tool should, therefore, be more widely applicable in a range clinical environments. Not all variables within the model would need to be modifiable for the model to be effective at identifying high-risk groups and allowing triage to effective management that might be independent of the specific risk factors identified. For example, within the widely used A-Age, B-Blood pressure, C-Clinical TIA features, D-Duration of symptoms, and D-Diabetes (ABCD2) score to identify those at high risk of stroke following transient ischemic attack,[77] only diabetes and blood pressure are modifiable. However, the tool is effective at helping prevent stroke because it identifies those at high risk effectively and therefore allows earlier implementation of stroke prevention strategies beyond simply controlling blood pressure and diabetes, including lipid profiling, cardiac monitoring, carotid imaging, and considering dual antiplatelets.[78]

Our model was a good fit for the data statistically, as demonstrated by a p-value above 0.05 on the Hosmer-Lemeshow goodness-of-fit test. The absolute figures for discrimination through sensitivity and specificity estimates were moderate, around 70%. This highlights some of the complexities of regression model interpretation, which measures predictive ability in several different ways. Discrimination is related to correctly separating individuals with and without disease. Therefore, it is more important in diagnostic accuracy studies, unlike ours. Calibration measures the agreement between observed and predicted risk, and is therefore more important in prognostic studies such as ours.[79] Calibration is demonstrated highly in our model by the predicted estimates mirroring the actual estimates closely throughout the Hosmer-Lemeshow contingency table (Table 5). This is reassuring for the clinical conclusions drawn from this model.

Finally, the study was designed to identify those at increased risk of epilepsy-related deaths as a group, but made no attempt to risk stratify according to the type of epilepsy-related death specifically, e.g. SUDEP, drowning, accidents or suicide. A future study could attempt to further stratify by type of epilepsy-related death following on from our demonstration of increased risks for the overall group.

18

1.5 Conclusion

We propose the SEDS Score as a potential tool for identifying adults at increased risk of epilepsy-related death based on their recent attendance history to hospital for seizures or epilepsy, area of deprivation, epilepsy aetiology profile, and their serious comorbidity burden. This 4-point score needs external validation in another healthcare system.[15,56] The SEDS Score has the potential to be applied clinically to help streamline frontline care for PWE and to inform clinical discussions with patients about risk. We describe additional risk factors to consider as part of the overall analysis of epilepsy-related mortality risk in PWE, including recent review by a neurologist, alcohol abuse, and mental health problems. There is the potential for a larger study to try to incorporate these into future iterations of the SEDS Score and for them to be included in other tools such as in the SUDEP and Seizure Safety Checklist.[11] The missing data patterns in our study for several other potentially modifiable risk factors for epilepsy-related indicate the need for better clinical documentation within medical records, perhaps with use of seizure history collection proformas in non-specialist areas. There is also a need for more granular administrative healthcare data sources to allow such variables to be studied at scale.

Acknowledgements

also grateful to Siddharthan Chandran, Catherine Sudlow, and William Whiteley for strategic support and advice on the project. We are grateful to Jane Andrews and Karen Chalmers for administerial support in relation to this work.

Funding

Competing interests

The authors report no competing interests.

Supplementary material

Supplementary material is available at *Brain* online.

References

1.   Singh A, Trevick S. The Epidemiology of Global Epilepsy. *Neurol Clin*. Nov 2016;34(4):837-847. doi:10.1016/j.ncl.2016.06.015

2.   Joint Epilepsy Council of the UK and Ireland. Epilepsy prevalence, incidence and other statistics. 2011;

3.   Mbizvo GK, Bennett K, Simpson CR, Duncan SE, Chin RFM. Epilepsy-related and other causes of mortality in people with epilepsy: A systematic review of systematic reviews. *Epilepsy Res*. Aug 27 2019;157:106192. doi:10.1016/j.eplepsyres.2019.106192

4.   Devinsky O, Spruill T, Thurman D, Friedman D. Recognizing and preventing epilepsy-related mortality: A call for action. *Neurology*. Feb 23 2016;86(8):779-86.

5. Mbizvo GK, Schnier C, Simpson CR, Chin RFM, Duncan SE. A national study of epilepsy-related deaths in Scotland: Trends, mechanisms, and avoidable deaths. *Epilepsia*. Nov 2021;62(11):2667-2684. doi:10.1111/epi.17065

6. Neligan A, Sander JW. Premature mortality in epilepsy: is it preventable? *Expert Review of Neurotherapeutics*. Jun 2011;11(6):767-770. doi:10.1586/Ern.11.77

7. Thurman DJ, Logroscino G, Beghi E, et al. The burden of premature mortality of epilepsy in high-income countries: A systematic review from the Mortality Task Force of the International League Against Epilepsy. *Epilepsia*. Jan 2017;58(1):17-26. doi:10.1111/epi.13604

8. Sander, Bell. Reducing mortality: an important aim of epilepsy management. *J Neurol Neurosurg Psychiatry*. Mar 2004;75(3):349-51.

9. Novak JL, Miller PR, Markovic D, Meymandi SK, DeGiorgio CM. Risk Assessment for Sudden Death in Epilepsy: The SUDEP-7 Inventory. *Front Neurol*. 2015;6:252. doi:10.3389/fneur.2015.00252

10. DeGiorgio CM, Miller P, Meymandi S, et al. RMSSD, a measure of vagus-mediated heart rate variability, is associated with risk factors for SUDEP: the SUDEP-7 Inventory. *Epilepsy Behav*. Sep 2010;19(1):78-81. doi:10.1016/j.yebeh.2010.06.011

11. SUDEP Action. SUDEP and Seizure Safety Checklist. Accessed 07/11/2018, https://bit.ly/347i8iY

12. Tarighati Rasekhi R, Devlin KN, Mass JA, et al. Improving prediction of sudden unexpected death in epilepsy: From SUDEP-7 to SUDEP-3. *Epilepsia*. Jun 4 2021;doi:10.1111/epi.16928

13. Walczak TS, Leppik IE, D'Amelio M, et al. Incidence and risk factors in sudden unexpected death in epilepsy: a prospective cohort study. *Neurology*. Feb 27 2001;56(4):519-25.

14. Jha A, Oh C, Hesdorffer D, et al. Sudden Unexpected Death in Epilepsy: A Personalized Prediction Tool. *Neurology*. May 25 2021;96(21):e2627-e2638. doi:10.1212/WNL.0000000000011849

21

15. Grant SW, Collins GS, Nashef SAM. Statistical Primer: developing and validating a risk prediction model. *Eur J Cardiothorac Surg*. Aug 1 2018;54(2):203-208. doi:10.1093/ejcts/ezy180

16. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. Jan 7 2015;350:g7594. doi:10.1136/bmj.g7594

17. Mbizvo GK, Schnier C, Simpson CR, Duncan SE, Chin RFM. Validating the accuracy of administrative healthcare data identifying epilepsy in deceased adults: A Scottish data linkage study. *Epilepsy Res*. Sep 13 2020;167:106462. doi:10.1016/j.eplepsyres.2020.106462

18. Chief Medical Officer. Guidance for Doctors Completing Medical Certificates of The Cause of Death (MCCD) and its Quality Assurance 2018. Accessed 03/03/2020. https://bit.ly/2yyMzlC

19. National Records of Scotland. Death Certificates and Coding the Causes of Death 2017. Accessed 31/07/2020. https://bit.ly/2DrrjR2

20. National Records of Scotland. Some Specific Checks Carried Out on Deaths Data. *Vital Events – Deaths – Background Information* 2019. Accessed 05/09/2020. https://bit.ly/2F50Gm6

21. NHS Education for Scotland. Medical certification of cause of death. 2021. Accessed 16/03/2022. https://bit.ly/3tZS8CO

22. Fernie CGM. Scotland already has world leading mortality review system. *Bmj-Brit Med J*. Feb 7 2019;364doi:10.1136/bmj.l395

23. McKinstry B, Sullivan FM, Vasishta S, et al. Cohort profile: the Scottish Research register SHARE. A register of people interested in research participation linked to NHS data sets. *BMJ Open*. Feb 1 2017;7(2):e013351. doi:10.1136/bmjopen-2016-013351

24. Sullivan F, McKinstry B, Vasishta S. The "All of Us" Research Program. *N Engl J Med*. Nov 7 2019;381(19):1883-1884. doi:10.1056/NEJMc1912496

25. Kee VR, Gilchrist B, Granner MA, Sarrazin NR, Carnahan RM. A systematic review of validated methods for identifying seizures, convulsions, or epilepsy using administrative and

claims data. *Pharmacoepidemiology and Drug Safety*. Jan 2012;21:183-193. doi:10.1002/pds.2329

26. 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*. Apr 2005;46(4):470-2. doi:10.1111/j.0013-9580.2005.66104.x

27. Fisher RS, Acevedo C, Arzimanoglou A, et al. ILAE official report: a practical clinical definition of epilepsy. *Epilepsia*. Apr 2014;55(4):475-82. doi:10.1111/epi.12550

28. Scottish Parliament Epilepsy Consortium Scotland. SUDEP Rates remain high in Scotland. *Cross Party Group Meeting*. 2019. Accessed 21/09/2020. http://www.epilepsyconsortiumscotland.co.uk/sudep-rates-scotland/

29. Fyfe I. Ean Virtual 2020 - the Largest Neurology Conference in History. *Nat Rev Neurol*. Aug 2020;16(8):404. doi:10.1038/s41582-020-0387-6

30. ERUK. ERUK Researcher Wins Award At ABN National Conference Epilepsy Research UK; 2019. Accessed 19/05/2021. https://bit.ly/3bDanWa

31. Information Services Division. The Scottish Index of Multiple Deprivation (SIMD). *GPD Support - Deprivation*. ISD Scotland; 2016. Accessed 27/10/2019. https://bit.ly/2V32ilM

32. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373-83.

33. Quan H, Li B, Couris CM, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *Am J Epidemiol*. Mar 15 2011;173(6):676-82. doi:10.1093/aje/kwq433

34. Radovanovic D, Seifert B, Urban P, et al. Validity of Charlson Comorbidity Index in patients hospitalised with acute coronary syndrome. Insights from the nationwide AMIS Plus registry 2002-2012. *Heart*. Feb 2014;100(4):288-94. doi:10.1136/heartjnl-2013-304588

35. Elovic A, Pourmand A. MDCalc Medical Calculator App Review. *J Digit Imaging*. Oct 2019;32(5):682-684. doi:10.1007/s10278-019-00218-y

23

36. Poupin P, Bouleti C, Degand B, et al. Prognostic value of Charlson Comorbidity Index in the elderly with a cardioverter defibrillator implantation. *Int J Cardiol*. Sep 1 2020;314:64-69. doi:10.1016/j.ijcard.2020.03.060

37. Kling R, Ritvanen J, Mustonen H, Kamppi L. Long-term outcome of convulsive status epilepticus: a 10-year follow-up. *Epileptic Disord*. Dec 1 2022;24(6):1-14. Long-term outcome of convulsive status epilepticus: a 10-year follow-up. doi:10.1684/epd.2022.1482

38. Tuty Kuswardhani RA, Henrina J, Pranata R, Anthonius Lim M, Lawrensia S, Suastika K. Charlson comorbidity index and a composite of poor outcomes in COVID-19 patients: A systematic review and meta-analysis. *Diabetes Metab Syndr*. Nov-Dec 2020;14(6):2103-2109. doi:10.1016/j.dsx.2020.10.022

39. eDRIS Team. Researcher guide: requesting outputs from Safe Haven and disclosure control 2018. Accessed 16/11/2019. https://bit.ly/2XzD1Bk

40. Bujang MA, Sa'at N, Sidik T, Joo LC. Sample Size Guidelines for Logistic Regression from Observational Studies with Large Population: Emphasis on the Accuracy Between Statistics and Parameters Based on Real Life Clinical Data. *Malays J Med Sci*. Jul 2018;25(4):122-130. doi:10.21315/mjms2018.25.4.12

41. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. *J Clin Epidemiol*. Dec 1995;48(12):1503-10. doi:10.1016/0895-4356(95)00048-8

42. Chowdhury MZI, Turin TC. Variable selection strategies and its importance in clinical prediction modelling. *Fam Med Community Health*. 2020;8(1):e000262. doi:10.1136/fmch-2019-000262

43. McNeish D. Missing data methods for arbitrary missingness with small samples. *J Appl Stat*. 2017;44(1):24-39. doi:10.1080/02664763.2016.1158246

44. Bennett DA. How can I deal with missing data in my study? *Aust Nz J Publ Heal*. Oct 2001;25(5):464-469.

45. IBM. Method (Multiple Imputation). *SPSS Statistics*. 2021. Accessed 19/05/2021. https://ibm.co/3wjcMNw

24

46. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. Wiley; 1987.

47. Dong YR, Peng CYJ. Principled missing data methods for researchers. *Springerplus*. 2013;2doi:Artn 222

10.1186/2193-1801-2-222

48. Goodman LA, Kruskal WH. *Measures of Association for Cross Classifications*. Springer-Verlag; 1979.

49. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J*. Aug 1 2014;35(29):1925-31. doi:10.1093/eurheartj/ehu207

50. Harrell FE. Multivariable Modeling Strategies. *Regression Modeling Strategies*. Springer Cham; 2015. Accessed 18/08/2022. https://link.springer.com/book/10.1007/978-3-319-19425-7

51. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. Springer; 2009.

52. Noe MH, Rosenbach M, Hubbard RA, et al. Development and Validation of a Risk Prediction Model for In-Hospital Mortality Among Patients With Stevens-Johnson Syndrome/Toxic Epidermal Necrolysis-ABCD-10. *JAMA Dermatol*. Apr 1 2019;155(4):448-454. doi:10.1001/jamadermatol.2018.5605

53. Dankers F, Traverso A, Wee L, van Kuijk SMJ. Prediction Modeling Methodology. In: Kubben P, Dumontier M, Dekker A, eds. *Fundamentals of Clinical Data Science*. 2019:101-120.

54. Muschelli J. ROC and AUC with a Binary Predictor: a Potentially Misleading Metric. *J Classif*. Dec 23 2019;doi:10.1007/s00357-019-09345-1

55. Mbizvo GK, Larner AJ. Receiver operating characteristic plot and area under the curve with binary classifiers: pragmatic analysis of cognitive screening instruments. *Neurodegener Dis Manag*. Oct 2021;11(5):353-360. doi:10.2217/nmt-2021-0013

56. Steyerberg EW, Moons KG, van der Windt DA, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med*. 2013;10(2):e1001381. doi:10.1371/journal.pmed.1001381

57. Dickson JM, Jacques R, Reuber M, et al. Emergency hospital care for adults with suspected seizures in the NHS in England 2007-2013: a cross-sectional study. *BMJ Open*. Oct 21 2018;8(10):e023352. doi:10.1136/bmjopen-2018-023352

58. Dixon PA, Kirkham JJ, Marson AG, Pearson MG. National Audit of Seizure management in Hospitals (NASH): results of the national audit of adult epilepsy in the UK. *BMJ Open*. Mar 31 2015;5(3):e007325. doi:10.1136/bmjopen-2014-007325

59. Mathieson A, Marson AG, Jackson M, et al. Clinically unnecessary and avoidable emergency health service use for epilepsy: A survey of what English services are doing to reduce it. *Seizure*. Feb 19 2020;76:156-160. doi:10.1016/j.seizure.2020.02.012

60. Ridsdale L, McCrone P, Morgan M, Goldstein L, Seed P, Noble A. *Can an epilepsy nurse specialist-led self-management intervention reduce attendance at emergency departments and promote well-being for people with severe epilepsy? A non-randomised trial with a nested qualitative phase*. 2013. *Health Services and Delivery Research*.

61. Noble AJ, Mathieson A, Ridsdale L, et al. Developing patient-centred, feasible alternative care for adult emergency department users with epilepsy: protocol for the mixed-methods observational 'Collaborate' project. *BMJ Open*. Nov 2 2019;9(11):e031696. doi:10.1136/bmjopen-2019-031696

62. Wojewodka G, Gulliford MC, Ashworth M, Richardson MP, Ridsdale L. Epilepsy and mortality: a retrospective cohort analysis with a nested case-control study identifying causes and risk factors from primary care and linkage-derived data. *BMJ Open*. Oct 25 2021;11(10):e052841. doi:10.1136/bmjopen-2021-052841

63. Mercado M, Gonzalez B, Vargas G, et al. Successful mortality reduction and control of comorbidities in patients with acromegaly followed at a highly specialized multidisciplinary clinic. *J Clin Endocrinol Metab*. Dec 2014;99(12):4438-46. doi:10.1210/jc.2014-2670

64. Levira F, Thurman DJ, Sander JW, et al. Premature mortality of epilepsy in low- and middle-income countries: A systematic review from the Mortality Task Force of the International League Against Epilepsy. *Epilepsia*. Dec 18 2016:1-11. doi:10.1111/epi.13603

65. Lyons J, Nafilyan V, Akbari A, et al. Validating the QCOVID risk prediction algorithm for risk of mortality from COVID-19 in the adult population in Wales, UK. *Int J Popul Data Sci*. 2020;5(4):1697. doi:10.23889/ijpds.v5i4.1697

66. Lyons RA, Jones KH, John G, et al. The SAIL databank: linking multiple health and social care datasets. *BMC Med Inform Decis Mak*. Jan 16 2009;9:3. doi:10.1186/1472-6947-9-3

67. Levy ML, Andrews R, Buckingham R, Evans H, Francis C, Houston R. Confidential Enquiry report. Why asthma still kills: The National Review of Asthma Deaths (NRAD) 2014. Accessed 19/08/2022. www.rcplondon.ac.uk/sites/default/files/why-asthma-still-kills-full-report.pdf

68. Wood A, Denholm R, Hollings S, et al. Linked electronic health records for research on a nationwide cohort of more than 54 million people in England: data resource. *Bmj-Brit Med J*. Apr 7 2021;372doi:ARTN n826

10.1136/bmj.n826

69. Buckley BJR, Harrison SL, Hill A, Underhill P, Lane DA, Lip GYH. Stroke-Heart Syndrome: Incidence and Clinical Outcomes of Cardiac Complications Following Stroke. *Stroke*. May 2022;53(5):1759-1763. doi:10.1161/STROKEAHA.121.037316

70. O'Neill B, Kalia S, Aliarzadeh B, et al. Cardiovascular risk factor documentation and management in primary care electronic medical records among people with schizophrenia in Ontario, Canada: retrospective cohort study. *BMJ Open*. Oct 16 2020;10(10):e038013. doi:10.1136/bmjopen-2020-038013

71. Franzen S, Janson C, Larsson K, et al. Evaluation of the use of Swedish integrated electronic health records and register health care data as support clinical trials in severe asthma: the PACEHR study. *Respir Res*. Nov 15 2016;17(1):152. doi:10.1186/s12931-016-0461-1

72. Haas AD, Ruffieux Y, van den Heuvel LL, et al. Excess mortality associated with mental illness in people living with HIV in Cape Town, South Africa: a cohort study using linked electronic health records. *Lancet Glob Health*. Oct 2020;8(10):e1326-e1334. doi:10.1016/S2214-109X(20)30279-5

73. Luczak A, Kalinowski S. Assessing the level of the material deprivation of European Union countries. *PLoS One*. 2020;15(9):e0238376. doi:10.1371/journal.pone.0238376

27

74. Kim AM, Rossi KC, Jette N, Yoo JY, Hung K, Dhamoon MS. Increased risk of hospital admission for mood disorders following admission for epilepsy. *Neurology*. Aug 28 2018;91(9):e800-e810. doi:10.1212/01.wnl.0000542492.00605.9d

75. Leung WCY, Lau EHY, Kwan P, Chang RS. Impact of COVID-19 on seizure-related emergency attendances and hospital admissions - A territory-wide observational study. *Epilepsy Behav*. Feb 2021;115:107497. doi:10.1016/j.yebeh.2020.107497

76. Navaratnam AV, Gray WK, Day J, Wendon J, Briggs TWR. Patient factors and temporal trends associated with COVID-19 in-hospital mortality in England: an observational study using administrative data. *Lancet Respir Med*. Apr 2021;9(4):397-406. doi:10.1016/S2213-2600(20)30579-8

77. Johnston SC, Rothwell PM, Nguyen-Huynh MN, et al. Validation and refinement of scores to predict very early stroke risk after transient ischaemic attack. *Lancet*. Jan 27 2007;369(9558):283-92. doi:10.1016/S0140-6736(07)60150-0

78. Pei LL, Chen P, Fang H, et al. Dual antiplatelet therapy reduced stroke risk in transient ischemic attack with positive diffusion weighted imaging. *Sci Rep*. Nov 5 2020;10(1):19132. doi:10.1038/s41598-020-75666-6

79. Kappen TH, van Klei WA, van Wolfswinkel L, Kalkman CJ, Vergouwe Y, Moons KGM. Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. *Diagn Progn Res*. 2018;2:11. doi:10.1186/s41512-018-0033-6

Figure legends

Figure 1 Study flow diagram. ASM – Antiseizure medication; *F25* – Primary care diagnostic Read codes for epilepsy; G40–41 – International Classification of Disease 10 (ICD-10) codes for epilepsy and status epilepticus; R56.8 – ICD-10 code for seizures; CHI – Scottish Community Health Index patient identification number; GP – General practitioner; NHS – National Health Service.

Table 1 Variables considered for analysis (N = 224 deceased cases and 224 alive controls)

Variable	Coding	Missing data burden		Chosen for multivariable model
		Cases	Controls	
SIMD Quintile[a]	Group 1 = 2 most deprived quintiles combined Group 2 = middle deprived quintile Group 3 = 2 least deprived quintiles combined	0%	3%	Yes
A&E attendance or hospital admission for epilepsy/seizures in preceding 12 months	Yes No Unable to tell (missing data)	4%	1%	Yes
Aetiology and risk factors for developing epilepsy	No risk factors for epilepsy Inherited/congenital[b] Acquired[c] Unable to tell (missing data)	13%	12%	Yes
Charlson Comorbidity Index[d]	0–1 2–3 ≥4 Unable to tell (missing data)	4%	13%	Yes
Alcohol abuse in preceding 24 months	Yes No Unable to tell (missing data)	11%	4%	No
Seizure type	Focal seizures ± secondary generalisation Generalised seizures	20%	9%	No
Mental health problems in preceding 24 months[e]	Yes No Unable to tell (missing data)	16%	14%	No
Epilepsy surgery ever	Yes No Unable to tell (missing data)	7%	1%	No
Seen by neurologist in preceding 12 months	Yes No Unable to tell (missing data)	1%	1%	No
Number of ASMs when last seen	Not on an ASM 1 ASM 2 ASMs ≥3 ASM Unable to tell (missing data)	0%	0%	No
Seizure frequency when last seen	<1 seizure per year ≥1 seizure per year Unable to tell (missing data)	40%	11%	No
Seizure timing	Mainly daytime Manly nocturnal Mixture of both daytime and nocturnal Unable to tell (missing data)	70%	56%	No
Seizure alarm in place	Yes No Unable to tell (missing data)	60%	36%	No
Status epilepticus in preceding 24 months	Yes No Unable to tell (missing data)	13%	1%	No

SIMD - Scottish Index of Multiple Deprivation; A&E – Accident and Emergency; ASMs – Antiseizure medications.

[a]Socioeconomic deprivation markers splitting the population into quintiles ordered 1 (most deprived) to 5 (least deprived) areas of residence.

[b]Inherited/congenital: Febrile convulsions, first degree relative with epilepsy, congenital abnormality/malformation (e.g. cerebral palsy, metabolic infancy syndrome, birth hypoxia), genetic syndrome, attention deficit hyperactivity disorder (ADHD), autism spectrum disorder (ASD), developmental/intellectual delay, premature birth, birth/perinatal difficulties, hydrocephalus, neonatal seizures.

[c]Acquired: History of meningitis/encephalitis, severe head injury, brain tumour, cerebrovascular disease, limbic encephalitis, brain surgery, CNS demyelinating disease, abnormal brain imaging, other hypoxic brain injury, neurodegenerative disease.

[d]Charlson Comorbidity Index: Predicts 10-year survival in patients with multiple comorbidities, accessible here: www.mdcalc.com/calc/3917/charlson-comorbidity-index-cci. Patients can have a score of 0–37, and we chose the initial grouping based on recent literature.[36]

[e]Includes psychosis, depression, anxiety, contact with psychiatric/psychological/learning disability services, admission under psychiatry, self-harm/suicide attempts.

1
2 **Table 2 Correlation results for variable pairs**

Variable	Variable	Lambda Correlation	
Seizure type	Aetiology and risk factors for developing epilepsy	0.31	Strong association
Seen by neurologist	Number of ASMs	0.28	Moderate association
Seen by neurologist	Charlson Comorbidity Index	0.20	Moderate association
Aetiology and risk factors for developing epilepsy	Charlson Comorbidity Index	0.19	Moderate association
Alcohol abuse	Aetiology and risk factors for developing epilepsy	0.15	Moderate association
A&E attendance or hospital admission	Mental health problems	0.12	Moderate association
SIMD Quintile	Aetiology and risk factors for developing epilepsy	0.07	Weak association
SIMD Quintile	Seen by neurologist	0.06	Weak association
Alcohol abuse	Mental health problems	0.05	Weak association

3 A&E – Accident and Emergency; ASMs – Antiseizure medications; SIMD – Scottish Index of Multiple Deprivation.
4
5

6 **Table 3 Univariable results**

Predictor variable	Grouping	Study status		Total	Complete case data analysis OR (95% CI)	P-value	Multiple imputation data OR (95% CI)	P-value
		Alive (control)	Deceased (case)					
SIMD Quintile	Two least deprived	85	50	135	Reference		Reference	
	Middle deprivation	42	43	85	1.8 (1.0–3.2)	0.06	1.8 (1.0–3.2)	0.05
	Two most deprived	91	131	222	**2.5 (1.6–4.0)**	**0.00**	**2.5 (1.6–4.0)**	**0.00**
Total		218	224	442				
Alcohol abuse	No	202	157	359	Reference		Reference	
	Yes	13	42	55	**4.4 (2.2–9.2)**	**0.000**	**3.7 (1.8–7.9)**	**0.001**
Total		215	199	414				
Seizure type	Focal epilepsy	153	119	272	Reference		Reference	
	Generalised epilepsy	71	105	176	**1.9 (1.2–3.0)**	**0.006**	**1.8 (1.1–2.8)**	**0.012**
Total		224	224	448				
A&E attendance or hospital admission	No	181	92	273	Reference		Reference	
	Yes	41	123	164	**5.1 (3.2–8.3)**	**0.000**	**5.1 (3.2–8.2)**	**0.000**
Total		222	215	437				
Mental health problems	No	101	80	181	Reference		Reference	
	Yes	91	109	200	**1.6 (1.0–2.6)**	**0.034**	**1.6 (1.1–2.4)**	**0.027**
Total		192	189	381				
Epilepsy surgery	No	193	203	396	Reference		Reference	
	Yes	29	6	35	**0.2 (0.1–0.5)**	**0.001**	0.4 (0.1–1.4)	0.159
Total		222	209	431				
Seen by neurologist (entire dataset)	Yes	180	115	295	Reference		Reference	
	No	42	106	148	**3.8 (2.4–6.1)**	**0.000**	**3.9 (2.5–6.3)**	**0.000**
Total		222	221	443				
Seen by neurologist (SHARE-only dataset)	**No**	26	61	87	**3.4 (1.8–6.1)**	**0.000**		
	Yes	96	61	157	Reference		Reference	
Total		122	122	244				
Aetiology and risk factors for developing epilepsy	None	92	55	147	Reference		Reference	
	Inherited/congenital	41	73	114	**3.1 (1.7–5.7)**	**0.000**	**3.3 (1.9–5.8)**	**0.000**
	Acquired	64	67	131	1.7 (0.9–3.2)	0.088	**1.9 (1.1–3.4)**	**0.030**

30

Total		197	195	392				
Charlson Comorbidity Index (CCI) score	0–1	143	127	270	Reference		Reference	
	2–3	33	61	94	**2.5 (1.2– 5.2)**	**0.010**	**3.3 (1.7–6.6)**	**0.000**
	≥4	19	27	46	2.0 (0.7–5.2)	0.200	**3.2 (1.2–8.3)**	**0.018**
Total		195	215	410				
Number of AEDs	Not on an ASM	3	8	11	Reference		Reference	
	1 ASM	101	96	197	0.4 (0.1–1.3)	0.127	0.3 (0.1–1.3)	0.336
	2 ASMs	67	75	142	0.4 (0.1–1.6)	0.201	0.4 (0.1–1.5)	0.397
	≥3 ASMs	53	44	97	0.3 (0.1–1.2)	0.098	0.3 (0.1–1.2)	0.289
Total		224	223	447				

1 SIMD – Scottish Index of Multiple Deprivation; A&E – Accident and Emergency; ASM – Antiseizure medication; CCI – Charlson Comorbidity
2 Index; SHARE - Scottish Health Research Register; OR – Odds ratio; CI – Confidence interval. Bold font within the table body corresponds to
3 statistically significant results.
4

5 **Table 4 Multivariable modelling**

Predictor variable	Grouping	Complete case data analysis OR (95% CI)	P-value	Bootstrap P-value	Multiple imputation data OR (95% CI)	P-value	SEDS Score weighting
SIMD Quintile	Two least deprived and two middle deprived	Reference			Reference		
	Two most deprived	**2.2 (1.2–3.8)**	**0.009**	**0.005**	**2.2 (1.3–3.8)**	**0.003**	**2**
A&E attendance or hospital admission	No	Reference			Reference		
	Yes	**4.2 (2.3–7.7)**	**0.000**	**0.001**	**4.7 (2.7–8.0)**	**0.000**	**3**
Aetiology and risk factors for developing epilepsy	No risk factors and Acquired risk factors	Reference			Reference		
	Inherited/congenital	**2.8 (1.4–5.6)**	**0.003**	**0.005**	**3.5 (1.8–6.5)**	**0.000**	**2**
CCI score	0–1	Reference			Reference		
	≥2	1.6 (0.6–4.3)	0.389	0.413	**2.6 (1.1–6.1)**	**0.034**	**1**

6 SIMD - Scottish Index of Multiple Deprivation; A&E – Accident and Emergency; CCI – Charlson Comorbidity Index; OR – Odds ratio; CI –
7 Confidence interval. Bold font within the table body corresponds to statistically significant results.
8

ACCEPTED MANUSCRIPT

Table 5 Contingency Table for Hosmer and Lemeshow Test

Multivariable model				SEDS Score model			
Study status = Alive (control)		Study status = Deceased (case)		Study status = Alive (control)		Study status = Deceased (case)	
Observed	Predicted	Observed	Predicted	Observed	Predicted	Observed	Predicted
35	35	6	6	59	59	12	12
33	35	17	15	27	27	10	10
21	19	8	10	102	102	73	73
18	17	9	10	29	29	81	81
21	19	19	21	7	7	48	48
10	13	23	20	–	–	–	–
9	9	23	23	–	–	–	–
8	9	28	27	–	–	–	–
<5	4	28	27	–	–	–	–
5	3	21	23	–	–	–	–

Table 6 SEDS Score model results

SEDS Score	Alive (control)	Deceased (case)	Complete case data analysis OR (95% CI)	P-value	Bootstrap P-value	Multiple imputation data OR (95% CI)	P-value
0	59	12	Reference			Reference	
1	27	10	1.6 (0.5–4.8)	0.443	0.456	1.7 (0.5–6.2)	0.412
2–3	102	73	**2.8 (1.3–6.2)**	**0.009**	**0.009**	**3.3 (1.3–8.5)**	**0.012**
4–5	29	81	**14.4 (5.9–35.2)**	**0.000**	**0.001**	**14.7 (5.4–39.9)**	**0.000**
6–8	7	48	**24.0 (8.1–71.2)**	**0.000**	**0.001**	**29.5 (9.1–96.3)**	**0.000**

OR – Odds ratio; CI – Confidence interval. Bold font within the table body corresponds to statistically significant results.
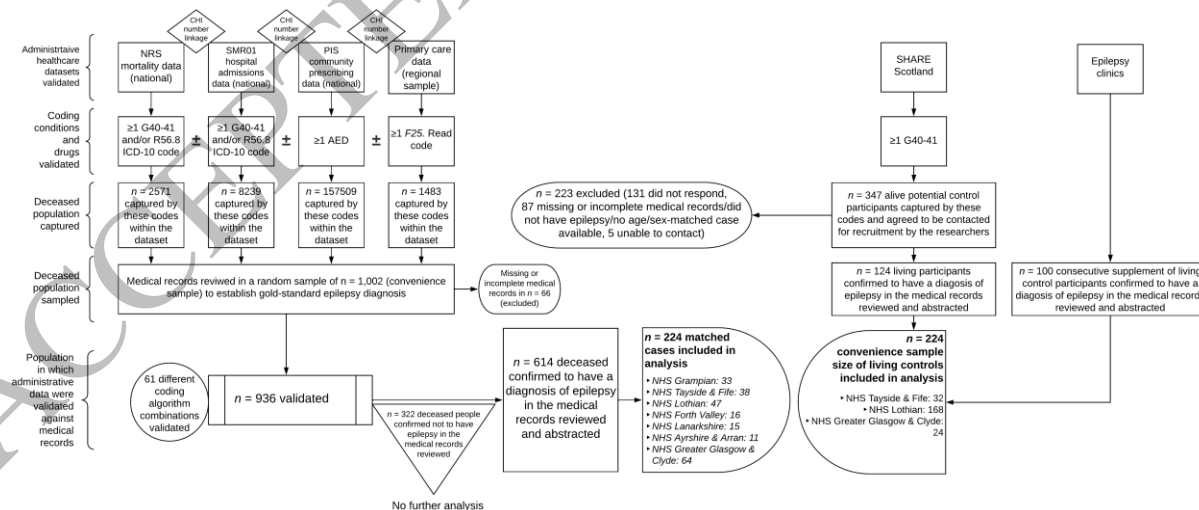

Figure 1
159x67 mm (3.8 x DPI)