National Institute for Health and Care Excellence

FINAL

Epilepsies in children, young people and adults: diagnosis and management

[1] Evidence review: Prediction of second seizure

NICE guideline NG217

Evidence reviews underpinning research recommendations in the NICE guideline

April 2022

FINAL

Developed by the National Guideline Centre

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or their carer or guardian.

Local commissioners and/or providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

Copyright

ISBN: 978-1-4731-4513-9

Contents

1. Prediction of a further seizure after a first seizure

1.1 Review question:

What are the most accurate tools for predicting a further seizure in people who have had a single seizure?

1.1.1. Introduction

A first seizure may significantly impact a person's life, including their social interactions, education, employment and driving privileges. The likelihood of having a further seizure is known to differ between individuals and between those with different underlying causes. Understanding and quantifying the likely risk of seizure recurrence will help people to manage its impact on their lives and inform their shared decision to start long term antiseizure medication. This review evaluates the accuracy of risk prediction tools for predicting who will go on to have a second seizure.

1.1.2. Summary of the protocol

For full details, see the review protocol in Appendix A.

Table 1: PICO characteristics of review question

Population	Inclusion: People with a history of a single seizure. These people are unlikely to have a fixed diagnosis of epilepsy. Exclusion: New-born babies with acute symptomatic seizures
Target condition	Epilepsy; suspected epilepsy
Prediction test	Any risk prediction tools for second seizure used clinically, performed at baseline.
Reference standard	Second seizure during subsequent follow up.
Statistical measures	Discrimination: sensitivity, specificity, C statistic. These measures assess how accurately the tool can predict those who will and will not have a second seizure. Calibration: tests how well the tool results predict the absolute risk of a second seizure. Net classification Improvement: a sensitive method for evaluating the different levels of predictive accuracy accruing from a change in the prediction tool. Follow up: use all available but stratify: <6 months, 6-12 months, 1-5 years, >5 years
Study design	Internal or external validation studies of the prediction tools. External validation studies (tested on a different study sample to the derivation sample) are preferred, although internal derivation studies (where the validation samples are different but still drawn from the identical population to the derivation sample) will still be included with a downgrade for indirectness. These validation studies will almost certainly be prospective cohort studies, but retrospective cohort studies will be used if available. In the context of validation studies, a retrospective design is less desirable as it permits a biased selection of cases that may yield results suggesting greater predictive accuracy.

1.1.3. Methods and process

This evidence review was developed using the methods and process described in Developing NICE guidelines: the manual. Methods specific to this review question are described in the review protocol in appendix A and the methods document.

Declarations of interest were recorded according to NICE's conflicts of interest policy.

1.1.4. Predictive evidence

1.1.4.1. Included studies

A search was made for studies that measure the accuracy of tools for predicting second seizures. Two studies were included in the review[3, 6]. These are summarised in Table 2 below, and details of the tools are summarised in Table 3. Evidence from these studies is summarised in the clinical evidence summaries below (Table 4 and Table 5) and Figure 1.

Studies were stratified according to two criteria [Age: young (<18) vs adults (≥18); Follow-up times: <6 months, 6-12 months, 1-5 years, >5 years] and thus placed in permutations of these categories. The included papers fitted into the following permuted strata: Adult/1-5 years follow up. Within this stratum, sub-grouping had been planned to try to 'explain' heterogeneity in meta-analyses according to the following strategies: young subgroups: <2, 2-11, 11-18; adults: 18-55, >55; learning disability vs no learning disability; head injury vs no head injury; type of epilepsy; gender. However, these sub-grouping strategies were not required because, in the absence of data that could be pooled, no heterogeneity could exist.

The assessment of the evidence quality was conducted with emphasis on discrimination measures such as sensitivity/specificity and the C statistic, as these were identified by the committee as the primary measures in guiding decision-making. The committee set clinical decision thresholds for:

- Sensitivity: 0.9 above which a test would be recommended and 0.6 below which a test is of no clinical use.
- Specificity: 0.5 above which a test would be recommended and 0.1 below which a test is of no clinical use.
- C statistics: 0.7 above which a test would be recommended and 0.5 below which a test is of no clinical use.

The committee placed greater emphasis on sensitivity than specificity because the harms of false negatives were agreed to be worse than the harms of false positives, in the context of second seizure prediction.

See also the study selection flow chart in Appendix C, and study evidence tables in Appendix D.

1.1.4.2. Excluded studies

See the excluded studies list in Appendix J.

.

1.1.5. Summary of studies included in the predictive evidence

Table 2: Summary of studies included in the evidence review

Study	Population	Predictive test	Reference standard (outcome event) definition	Number of outcome events	Follow up duration
Bonnett, 2014[3]	N=1426; 3 separate cohorts of adults with a first seizure evaluated separately. Median ages of cohorts varied between 28 and 50.3. Taken from secondary care patient data in UK, Italy and Australia. Type of epilepsy was unclear.	Model developed by the MESS trial (Bonnett, 2010), which aims to predict recurrent seizure within 12 months of the first seizure. The model includes variables for aetiology, epilepsy in a first-degree relative, seizure while asleep, electroencephalogram (EEG) results, CT or MRI imaging and treatment policy. Risk groups were determined by the 16th, 50th and 84th centiles – thus low risk (0-16), moderate risk (17-50), high risk (51-84) and very high risk (85-100).	Second seizure (not defined)	Not reported	12 months
Kim, 2016[6]	N=124. Cohort of adults from South Korea with a single seizure post-ischaemic stroke. Median age 68; 44% female; 54.8% generalised seizures, 43.2% partial seizures.	Three tools were evaluated: 1. The study prediction tool comprised 10 predictors: seizure onset under 65yrs, male gender, AF, lesion size, cortex involvement, haemorrhagic transformation, functional disability after stroke, status epilepticus, stroke lesion EEG findings, partial seizure. 7 versions of these were tested (1-1, 1-2, 2-1, 2-2, 3-1, 3-2, 4; these are defined in the table below) 2. Post-Stroke Epilepsy Risk Scale (PoSERS) tool: supratentorial stroke, cortical involvement, haemorrhagic transformation, modified Rankin score >3	Unprovoked second seizure separated from first one by >24hrs. Two sub-classes: <7-day recurrence (early PSSI) and ≥7-day recurrence (late PSSI)	54	Mean 29.9 months post first seizure

Study	Population	Predictive test	Reference standard (outcome event) definition	Number of outcome events	Follow up duration
		3. MESS tool: modified Rankin score >1, abnormal EEG			

Table 3: Summary of prediction tools used in the included studies and constituent variables and cut-offs (where available)

Risk tool	Variables and scoring
Model developed by the MESS trial (Bonnett, 2010), which aims to predict recurrent seizure within 12 months of first seizure.	The model includes variables for aetiology, epilepsy in a first-degree relative, seizure while asleep, electroencephalogram (EEG) results, CT or MRI imaging and treatment policy. Risk groups were determined by the 16th, 50th and 84th centiles – thus low risk (0-16), moderate risk (17-50), high risk (51-84) and very high risk (85-100).
Kim, 2016 prediction tool version 1-1	The study prediction tool comprised 10 predictors: seizure onset under 65yrs [1 point], male gender [1 point], AF [1 point], lesion size [small = 0 points, moderate=1 point, large = 2points], cortical involvement [1 point], haemorrhagic transformation [1 point], functional disability after stroke [mild= 0 point, moderate=1 point, severe=2 point], status epilepticus [1 point], relevant focal EEG findings[1 point], partial seizure[1 point]. Threshold = 6 for 'early PSSI' (<7 days between stroke and first seizure) and 7 for 'late PSSi' (>7 days between stroke and first seizure).
Kim, 2016 prediction tool version 1-2	The study prediction tool comprised 10 predictors: seizure onset under 65yrs [1 point], male gender [1 point], AF [1 point], lesion size [small = 0 points, moderate=1 point, large = 2points], cortical involvement [1 point], haemorrhagic transformation [1 point], modified Rankin scale, status epilepticus [1 point], relevant focal EEG findings[1 point], partial seizure[1 point]. Threshold = 9
Kim, 2016 prediction tool version 2-1	The study prediction tool comprised 7 predictors: lesion size [small = 0 points, moderate=1 point, large = 2points], cortical involvement [1 point], haemorrhagic transformation [1 point], functional disability after stroke [mild= 0 point, moderate=1 point, severe=2 point], status epilepticus [1 point], relevant focal EEG findings[1 point], partial seizure[1 point]. Threshold = 5

Risk tool	Variables and scoring
Kim, 2016 prediction tool version 2-2	The study prediction tool comprised 7 predictors: lesion size [small = 0 points, moderate=1 point, large = 2points], cortical involvement [1 point], haemorrhagic transformation [1 point], modified Rankin scale, status epilepticus [1 point], relevant focal EEG findings[1 point], partial seizure[1 point]. Threshold = 8 for early PSSI (<7 days between stroke and first seizure) and 7 for 'late PSSi' (>7 days between stroke and first seizure).
Kim, 2016 prediction tool version 3-1	The study prediction tool comprised 5 predictors: male gender [1 point], AF [1 point], cortical involvement [1 point], functional disability after stroke [mild= 0 point, moderate=1 point, severe=2 point], partial seizure[1 point]. Threshold = 3
Kim, 2016 prediction tool version 3-2	The study prediction tool comprised 5 predictors: male gender [1 point], AF [1 point], cortical involvement [1 point], modified Rankin scale, partial seizure[1 point]. Threshold = 6
Kim, 2016 prediction tool version 4	The study prediction tool comprised 4 predictors: seizure onset under 65yrs [1 point], male gender [1 point], AF [1 point], lesion size [small = 0 points, moderate=1 point, large = 2points], partial seizure[1 point]. Threshold = 3
Post-Stroke Epilepsy Risk Scale (PoSERS) tool [9]	Supratentorial stroke, cortical involvement, haemorrhagic transformation, modified Rankin score >3. Reference to the original derivation paper[9] does not provide details of the scoring nor confirmation of the variables outlined in Kim, 2016 [6]
MESS tool in the study by Kim, 2016[6]: this appears to be very different to that described in Bonnett, 2014.	Modified Rankin score >1, abnormal EEG

(a) See Appendix D for full evidence tables

1.1.6. Summary of the predictive evidence

Adult/1-5 year follow up stratum

Discrimination

Table 4: Clinical evidence profile: Discriminative capacity (C statistic) of prediction tools featured in the studies (see table 3).

Prediction tool	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Area Under Curve Individual study effects [point estimate (95% Cis)] Pooled effect/range /median	Quality
MESS tool (As reported by Bonnett, 2014[3]	1	274(NGPSE cohort)	serious risk of bias[a]	NA	No serious indirectness	No serious imprecision	NGPSE: 0.60(0.55-0.65)	MODERATE
MESS tool (As reported by Bonnett, 2014[3]	1	847(WA cohort)	serious risk of bias[a]	NA	No serious indirectness	No serious imprecision	WA: 0.59(0.56-0.62)	MODERATE
MESS tool (As reported by Bonnett, 2014[3]	1	305 (FIRST Cohort)	serious risk of bias[a]	NA	No serious indirectness	No serious imprecision	FIRST: 0.65 (0.60-0.70). This was using multiple imputation to adjust for the fact that the sleep variable was missing. Other methods of imputation gave very similar results.	MODERATE

Prediction tool	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Area Under Curve Individual study effects [point estimate (95% Cis)] Pooled effect/range /median	Quality
Kim, 2016 prediction tool version 1-1	1	48	serious risk of bias[a]	NA	No serious indirectness	Serious imprecision[b]	Early PSSi: 0.653(0.502-0.784)	LOW
Kim, 2016 prediction tool version 1-1	1	76	serious risk of bias[a]	NA	No serious indirectness	Serious imprecision[b]	Late PSSi: 0.566(0.447-0.679)	LOW
Kim, 2016 prediction tool version 1-2	1	48	serious risk of bias[a]	NA	No serious indirectness	Very serious imprecision[b]	Early PSSi: 0.650(0.499-0.782)	VERY LOW
Kim, 2016 prediction tool version 1-2	1	76	serious risk of bias[a]	NA	No serious indirectness	Serious imprecision[b]	Late PSSi: 0.558(0.439-0.671)	LOW

Prediction tool	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Area Under Curve Individual study effects [point estimate (95% Cis)] Pooled effect/range /median	Quality
Kim, 2016 prediction tool version 2-1	1	48	serious risk of bias[a]	NA	No serious indirectness	Very serious imprecision[b]	Early PSSi: 0.621(0.470-0.757)	VERY LOW
Kim, 2016 prediction tool version 2-1	1	76	serious risk of bias[a]	NA	No serious indirectness	Serious imprecision[b]	Late PSSi: 0.534(0.416-0.649)	LOW
Kim, 2016 prediction tool version 2-2	1	48	serious risk of bias[a]	NA	No serious indirectness	Very serious imprecision[b]	Early PSSi: 0.622(0.471-0.758)	VERY LOW
Kim, 2016 prediction tool version 2-2	1	76	serious risk of bias[a]	NA	No serious indirectness	Serious imprecision[b]	Late PSSi: 0.526(0.408-0.642)	LOW
Kim, 2016 prediction	1	48	serious risk of bias[a]	NA	No serious indirectness	Serious imprecision[b]	Early PSSi: 0.735(0.588-0.852)	LOW

Prediction tool	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Area Under Curve Individual study effects [point estimate (95% Cis)] Pooled effect/range /median	Quality
tool version 3-1								
Kim, 2016 prediction tool version 3-2	1	48	serious risk of bias[a]	NA	No serious indirectness	Serious imprecision[b]	Early PSSi: 0.676(0.525-0.803)	LOW
Kim, 2016 prediction tool version 4	1	76	serious risk of bias[a]	NA	No serious indirectness	Serious imprecision[b]	Late PSSi: 0.734(0.620-0.829)	LOW
Post-Stroke Epilepsy Risk Scale (PoSERS) tool [9]	1	48	serious risk of bias[a]	NA	No serious indirectness	Very serious imprecision[b]	Early PSSi: 0.576(0.425-0.717)	VERY LOW
Post-Stroke Epilepsy Risk Scale (PoSERS) tool [9]	1	76	serious risk of bias[a]	NA	No serious indirectness	Serious imprecision[b]	Late PSSi: 0.532(0.414-0.647)	LOW

Prediction tool	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Area Under Curve Individual study effects [point estimate (95% Cis)] Pooled effect/range /median	Quality
MESS tool in the study by Kim, 2016[6]: this appears to be very different to that described in Bonnett, 2014.	1	48	serious risk of bias[a]	NA	No serious indirectness	Serious imprecision[b]	Early PSSi: 0.509(0.361-0.657	LOW
MESS tool in the study by Kim, 2016[6]: this appears to be very different to that described in Bonnett, 2014.	1	76	serious risk of bias[a]	NA	No serious indirectness	Very serious imprecision[b]	Late PSSi: 0.594(0.475-0.705)	VERY LOW

a) *Risk of bias was assessed using the PROBAST checklist (see Appendix F).Risk of bias was serious for all risk tools because none of the studies reported any blinding of assessors for risk tool data and outcome status.*
b) *The judgement of precision was based on the spread of confidence interval across two clinical thresholds: C statistics of 0.5 and 0.7. The threshold of 0.5 marked the boundary between no predictive value better than chance and a predictive value better than chance. The threshold of 0.7 marked the boundary above which the committee might consider recommendations. If the 95% Cis crossed one of these thresholds a rating of serious imprecision was given and if they crossed both of these thresholds a rating of very serious imprecision as given.*

Table 5: Clinical evidence profile: sensitivity and specificity of prediction tools featured in the studies (see table 3).

Prediction tool	No of studies	n	Sensitivity	Specificity	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Kim, 2016 prediction tool version 1-1 – EARLY PSSi (threshold* = 6 points)	1	48	**0.588**	**0.677**	**Sensitivity**				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
					specificity				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
Kim, 2016 prediction tool version 1-1 – LATE PSSi (threshold = 7 points)	1	76	**0.460**	**0.615**	**Sensitivity**				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
					specificity				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
Kim, 2016 prediction tool version 1-2 – EARLY PSSi	1	48	**0.529**	**0.807**	**Sensitivity**				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW

Prediction tool	No of studies	n	Sensitivity	Specificity	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
(threshold = 9 points)					**specificity**				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
Kim, 2016 prediction tool version 1-2 – LATE PSSi (threshold = 9 points)	1	76	**0.405**	**0.692**	**Sensitivity**				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
					specificity				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
Kim, 2016 prediction tool version 2-1 – EARLY PSSi (threshold = 5 points)	1	48	**0.647**	**0.710**	**Sensitivity**				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
					specificity				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
Kim, 2016 prediction tool version	1	76	**0.432**	**0.590**	**Sensitivity**				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor	LOW

Prediction tool	No of studies	n	Sensitivity	Specificity	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
2-1 – LATE PSSi (threshold = 5 points)								raw data provided[b]	
					specificity				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
Kim, 2016 prediction tool version 2-2 – EARLY PSSi (threshold = 8 points)	1	48	**0.529**	**0.807**	**Sensitivity**				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
					specificity				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
Kim, 2016 prediction tool version 2-2 – LATE PSSi (threshold = 7 points)	1	76	**0.541**	**0.590**	**Sensitivity**				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
					specificity				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
Kim, 2016 prediction	1	48	**0.706**	**0.710**	**Sensitivity**				

Prediction tool	No of studies	n	Sensitivity	Specificity	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
tool version 3-1 – EARLY PSSi (threshold = 3 points)					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
					specificity				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
Kim, 2016 prediction tool version 3-2 – EARLY PSSi (threshold = 6 points)	1	48	**0.471**	**0.807**	**Sensitivity**				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
					specificity				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
	1	76	**0.622**	**0.811**	**Sensitivity**				

Prediction tool	No of studies	n	Sensitivity	Specificity	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Kim, 2016 prediction tool version 4 – LATE PSSi (threshold = 3 points)					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
					specificity				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
MESS tool in the study by Kim, 2016[6]: this appears to be very different to that described in Bonnett, 2014): EARLY PSSi (threshold = 1 points)	1	48	**0.706**	**0.355**	**Sensitivity**				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
					specificity				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
MESS tool in the study by Kim, 2016[6]: this appears to be very	1	76	**0.324**	**0.872**	**Sensitivity**				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW

Prediction tool	No of studies	n	Sensitivity	Specificity	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
different to that described in Bonnett, 2014): LATE PSSi (threshold = 3 points)					**specificity**				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
PoSERS: EARLY PSSi (threshold = 3 points)	1	48	**0.588**	**0.581**	**Sensitivity**				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
					specificity				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
PoSERS: LATE PSSi (threshold = 4 points)	1	76	**0.216**	**0.897**	**Sensitivity**				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW
					specificity				
					Serious risk of bias[a]	NA	No serious indirectness	NA – no 95% CIs nor raw data provided[b]	LOW

a) Risk of bias was assessed using the PROBAST checklist. Risk of bias was serious for all risk tools because none of the studies reported any blinding of assessors for risk tool data.

b) No 95% CIs available so an assumption of serious imprecision has been made

** This number was termed 'criteria' in the paper but appears to relate to the threshold at which risk of second seizure would be regarded as higher. It is unclear whether this means that the upper-risk category would be at a value **more than** the value or **more than or equal** to the value. The chosen threshold appears to have been determined from the ROC curves and was not pre-specified. Thus, there may be some overestimation of accuracy through an 'artificial selection' mechanism (chance effects leading to high sensitivity or specificity values are selected for as the 'best threshold')*

Calibration

The charts below have been pasted directly from Bonnett, 2014 [3]. The charts represent how well the predictions made by the MESS tool agree with the observed data on second seizures from each validation cohort. The 'First' cohort was analysed with a variety of imputation strategies to account for missing data, but all strategies yielded very similar results. No quantitative data were provided by the paper.

Figure 1: Calibration plots for the MESS tool in the NGPSE, WA and FIRST cohorts

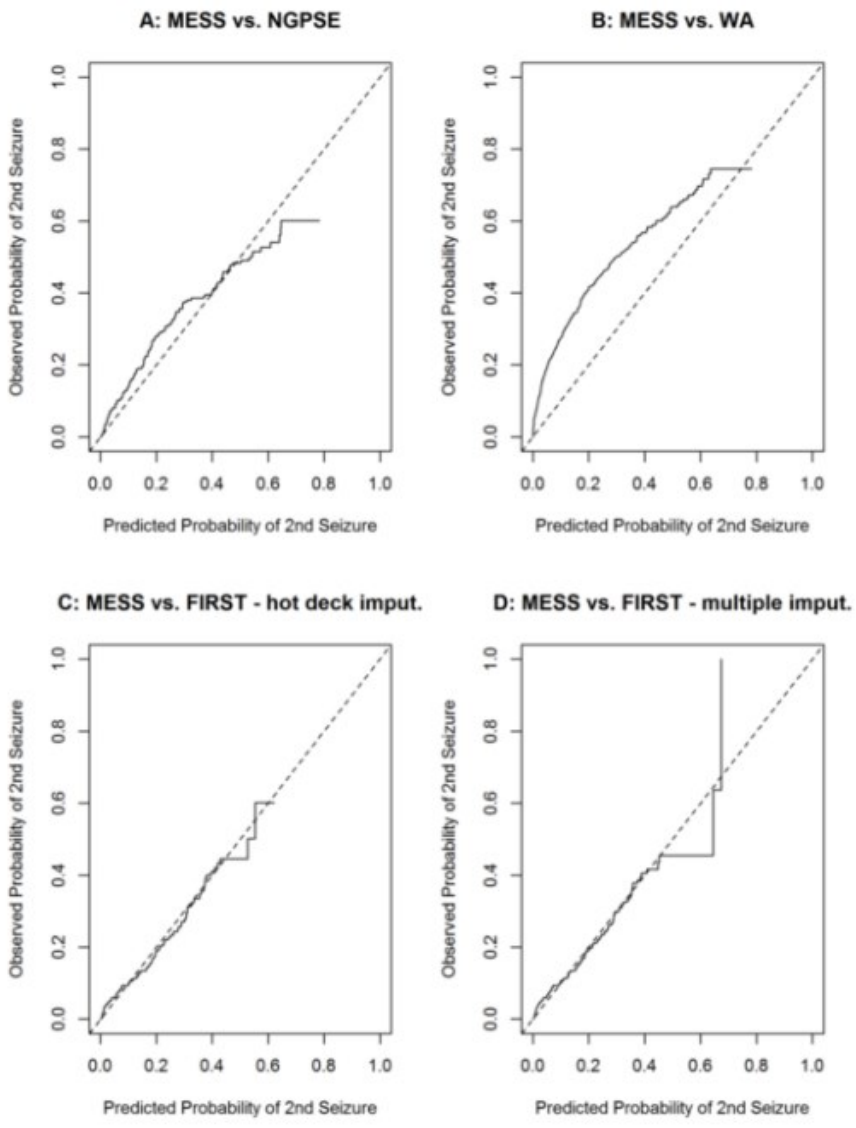


Figure 2. Calibration plots for MESS compared to NGPSE (A), WA (B), FIRST - variable matching (C) and FIRST - hot deck imputation (D).
doi:10.1371/journal.pone.0099063.g002

1.1.7. Economic evidence

1.1.7.1. Included studies

No health economic studies were included.

1.1.7.2. Excluded studies

No relevant health economic studies were excluded due to assessment of limited applicability or methodological limitations.

See also the health economic study selection flow chart in Appendix G.

1.1.8. Economic model

This area was not prioritised for a new cost-effectiveness analysis.

1.1.9. Evidence statements

1.1.9.1. Clinical evidence statements

- No relevant published evidence was identified.

1.1.9.2. Economic

- No relevant economic evaluations were identified.

1.1.10. The committee's discussion and interpretation of the evidence

1.1.10.1. The outcomes that matter most

The committee agreed that sensitivity and specificity were critical outcomes. It is vital to know how many people who go on to have a second seizure will be incorrectly labelled as low risk by the prediction tool (low sensitivity means more false negatives). It is also important to know how many people who will not go on to have a second seizure will be mistakenly labelled as high risk by the prediction tool (low specificity means a higher number of false positives). Knowledge of the likelihood of false negatives and false positives is essential so that clinicians can use the tools, ensuring that 1) patients at high risk will not be missed, and 2) patients at low risk will not be given inappropriately high levels of surveillance and anxiety. Sensitivity was deemed to be more important than specificity because the harms of false negatives are worse than the harms of false positives in the context of second seizure prediction. This arises because a false negative result will lead to patients who require prophylactic management not receiving the care that they need, which may cause harm. In contrast, a false positive result may lead to increased costs and anxiety, but is unlikely to lead to physically dangerous sequelae. However, specificity still needs to be at least moderate, as 100% sensitivity with very poor specificity is no better than simply assuming that every patient with a first seizure is at high risk of a second seizure.

C statistics were regarded as less important because they do not differentiate between sensitivity and specificity (from which they are derived), even though sensitivity may be more important in this context.

Calibration statistics were regarded as of equal status to sensitivity, as they allow an accurate evaluation of the agreement between the absolute risks yielded by the tools and the observed risks at all levels of risk; accurate risk evaluation may be of great importance when discussing results with the patient.

1.1.10.2. The quality of the evidence

The findings by Bonnett (2014) were of moderate quality, with a single downgrade because of the lack of blinding for outcome and prediction tool assessors. The findings by Kim (2016) were affected by the same issue of no blinding but also affected by imprecision, as the confidence intervals crossed one or both of the minimum important differences. This poor imprecision is likely to have been secondary to a very small sample size. In addition, Kim (2016) was an internal validation study, and it is unclear if similar effects would have been observed in an external sample. Findings were thus graded as low to very low.

1.1.10.3. Benefits and harms

The committee agreed that none of the reviewed tools were accurate enough to be used to predict who may go on to get a second seizure. The MESS tool, as evaluated by Burnett (2014), had good calibration in an Italian cohort but tended to underestimate risk in the UK and particularly Australian cohorts. This indicated that in UK and Australian populations, the MESS tool may tend to fail to detect some of those that will have a second seizure. Moreover, the C statistics data showed that the tool's discrimination ability was not up to the level (area under the ROC curve = 0.7) that would be regarded as 'good'. Furthermore, sensitivity and specificity data for each threshold of the tool were not presented in the paper, preventing appraisal of whether levels of sensitivity and specificity were adequate.

The 7 different Kim (2016) prediction tool versions, as well as the PoSERS tool and their version of the MESS tool, had similarly low C statistics to those presented by Burnett (2014), with reported sensitivities and specificities that did not suggest they had clinical utility. Moreover, whilst the analysis by Burnett (2014) were in cohorts that were representative of the vast majority of people with second seizure, the analyses by Kim (2016) were in a post-stroke population, which is a small sub-section of the total first seizure population. In summary, the committee's main concern with the evidence from both studies was that it suggested that people destined to have a second seizure might often fail to be predicted by the tools, leading to a lack of appropriate pre-emptive care.

The committee noted that as it is not common practice to use any prediction tool at all, recommending a tool such as the MESS tool could lead to some people at high risk of a second seizure being correctly detected, which would be an improvement on current practice. Because sensitivity would effectively move from zero (with no tool) to 'some' sensitivity (with a sub-optimal tool). However, the committee noted this strategy (of using a tool in place of no previous tool) would also lead to an increase in false positives and the harms of false positives – that is, incorrectly predicting high risk when the actual risk is low - were highlighted. Even though the harms of false positives may be less serious than the harms of failing to detect someone who will subsequently have a second seizure, they are not negligible and need to be considered. The committee was aware of other areas such as the use of prostate specific antigen test, where the sub-optimal specificity leads to great uncertainty within patients, with false positives causing considerable patient anxiety. In addition, it was argued that the time required to use a prediction tool would need to be reimbursed by adequate predictive accuracy, which was not deemed to be the case with the tools reviewed. The committee also acknowledged that experienced clinicians do use ad-hoc testing and clinical judgement to predict further seizures.

One concern with <u>not</u> recommending a prediction tool was that this might reduce clinician confidence in the measurement of each tools' constituent variables, such as EEG. This concern was discussed, but the committee's conclusion was that although the measurement of individual variables might not provide adequate predictive value, there is still a need for such measurements on grounds other than risk prediction. For example, MRI is important for excluding brain tumours. Therefore, the committee thought that not recommending a prediction tool would not dissuade clinicians from carrying out important clinical tests.

Reasons for the relatively disappointing predictive accuracy of the tools were discussed. One reason was the possibility of inferior technical specifications of measurements such as EEG or imaging at the time of the Burnett studies. In addition, the Burnett studies did not include every first seizure, as some people had their second seizure very quickly, prohibiting many from this sub-group from being eligible for the developmental and validation studies. Thus, the studies may not have been able to model the risk factors for early second seizures.

The committee discussed the great importance of being able to predict a second seizure in order to initiate early management and target people for whom reduction of modifiable risk factors might be beneficial. Current practice is to use clinical judgement about which clinical tests to use and how to interpret them. Although this process was deemed to have predictive accuracy that might be similar to the sub-optimal tools reviewed, it was still not regarded by the committee as sufficient basis upon which to base clinical practice, and it was agreed that far better prediction tools were required. The committee did not recommend any of the reviewed prediction tools for second seizure because they were considered to carry the potential for patient harm. This harm is a consequence of the tools' poor capacity to discriminate between people at low and high risk of second seizure, manifested by an unacceptably high proportion of false-negative and false-positive findings reported, both of which have the potential for harm. Therefore, the committee agreed a research recommendation should be made to develop and test a prediction tool for second seizure.

1.1.10.4. Cost effectiveness and resource use

No economic evidence was identified for this review.

The committee made a research recommendation; therefore, this will not result in substantial resource impact.

1.1.11. Recommendations supported by this evidence review

No recommendations were made from this evidence review. This evidence review supports the research recommendations on a prediction tool developmental study and an external validation study that involves testing the risk prediction tool(s).

References

1. Anonymous. Erratum: External validation of a prognostic model for seizure recurrence following a first unprovoked seizure and implications for driving (PLoS ONE (2014) 9:6 (e99063) DOI: 10.1371/journal.pone.0099063). PloS One. 2014; 9(11):e113327

2. Berg AT, Shinnar S. The risk of seizure recurrence following a first unprovoked seizure: a quantitative review. Neurology. 1991; 41(7):965-972

3. Bonnett LJ, Marson AG, Johnson A, Kim L, Sander JW, Lawn N et al. External validation of a prognostic model for seizure recurrence following a first unprovoked seizure and implications for driving. PloS One. 2014; 9(6):e99063

4. Bonnett LJ, Tudur-Smith C, Williamson PR, Marson AG. Risk of recurrence after a first seizure and implications for driving: further analysis of the Multicentre study of early Epilepsy and Single Seizures. BMJ. 2010; 341:c6477

5. Choquet C, Depret-Vassal J, Doumenc B, Sarnel S, Casalino E. Predictors of early seizure recurrence in patients admitted for seizures in the Emergency Department. European Journal of Emergency Medicine. 2008; 15(5):261-267

6. Kim HJ, Park KD, Choi KG, Lee HW. Clinical predictors of seizure recurrence after the first post-ischemic stroke seizure. BMC Neurology. 2016; 16(1):212

7. Kim LG, Johnson TL, Marson AG, Chadwick DW, group MMS. Prediction of risk of seizure recurrence after a single seizure and early epilepsy: further results from the MESS trial. Lancet Neurology. 2006; 5(4):317-322

8. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual [updated October 2020]. London. National Institute for Health and Care Excellence, 2014. Available from: http://www.nice.org.uk/article/PMG20/chapter/1%20Introduction%20and%20overview

9. Strzelczyk A, Haag A, Raupach H, Herrendorf G, Hamer HM, Rosenow F. Prospective evaluation of a post-stroke epilepsy risk scale. Journal of Neurology. 2010; 257(8):1322-1326

10. Tews W, Weise S, Syrbe S, Hirsch W, Viehweger A, Merkenschlager A et al. Is there a predictive value of EEG and MRI after a first afebrile seizure in children? Klinische Padiatrie. 2015; 227(2):84-88

Appendices

Appendix A Review protocols

A.1 Review protocol for prediction of second seizure

ID	Field	Content
0.	PROSPERO registration number	Not registered
1.	Review title	Prediction of a further seizure after a first seizure.
2.	Review question	What are the most accurate tools for predicting a further seizure, in people who have had a single seizure?
3.	Objective	To evaluate the best risk prediction tools for predicting who will go on to have a second seizure. Being able to predict who may have a second seizure is important because it may allow earlier management in those deemed at highest risk, which may improve long term outcomes.
4.	Searches	**The following databases (from inception) will be searched:** • Cochrane Central Register of Controlled Trials (CENTRAL) • Cochrane Database of Systematic Reviews (CDSR) • Embase • MEDLINE **Searches will be restricted by:** • English language **Other searches:** • None The searches may be re-run 6 weeks before the final committee meeting and further studies retrieved for inclusion if relevant. The full search will be published in the final review.
5.	Condition or domain being studied	Epilepsies
6.	Population	Inclusion: People with a history of a single seizure. These people are unlikely to have a fixed diagnosis of epilepsy. Exclusion: New-born babies with acute symptomatic seizures

7.	Predictor	Any risk prediction tools for second seizure used clinically.
8.	Types of study to be included	Internal or external validation studies of the prediction tools. External validation studies (tested on a different study sample to the derivation sample) are preferred, although internal derivation studies (where the validation sample are different, but still drawn from the identical population to the derivation sample) will still be included with a downgrade for indirectness. These validation studies will almost certainly be prospective cohort studies, but retrospective cohort studies will be used if available.
9.	Other exclusion criteria	Case-control studies, cross-sectional studies Non-English-language studies
10.	Context	There is evidence that second seizure (which often confirms epilepsy diagnosis) may be preventable in some people, and it is therefore important to be able to predict who is likely to develop a second seizure so that preventative actions (such as risk modification and earlier onset of management) can be affected.
11.	Primary outcomes (critical outcomes)	Discrimination: sensitivity, specificity, C statistic. These measures assess how accurately the tool can predict those who will and will not, have a second seizure. Calibration: tests how well the tool results predict the absolute risk of a second seizure. Net classification Improvement: a sensitive method for evaluating the different levels of predictive accuracy accruing from a change in the prediction tool. Follow up: use all available but stratify: <6 months, 6-12 months, 1-5 years, >5 years
12.	Secondary outcomes (important outcomes)	None
13.	Data extraction (selection and coding)	EndNote will be used for reference management, sifting, citations and bibliographies. All references identified by the searches and from other sources will be screened for inclusion. 10% of the abstracts will be reviewed by two reviewers, with any disagreements resolved by discussion or, if necessary, a third independent reviewer. The full text of these potentially eligible studies will be retrieved and assessed in line with the criteria outlined above. A standardised form will be used to extract data from the included studies (see Developing NICE guidelines: the manual section 6.4). 10% of all evidence reviews are quality assured by a senior research fellow. This includes checking: • papers were included /excluded appropriately • a sample of the data extractions • correct methods are used to synthesise data • a sample of the risk of bias assessments Disagreements between the review authors over the risk of bias in particular studies will be resolved by discussion, with involvement of a third review author where necessary.

14.	Risk of bias (quality) assessment	Risk of bias quality assessment will be assessed using PROBAST. 10% of all evidence reviews are quality assured by a senior research fellow. This includes checking: • papers were included /excluded appropriately • a sample of the data extractions • correct methods are used to synthesise data • a sample of the risk of bias assessments Disagreements between the review authors over the risk of bias in particular studies will be resolved by discussion, with involvement of a third review author where necessary.
15.	Strategy for data synthesis	Where possible suitably adjusted data will be meta-analysed where appropriate. Sensitivity and specificity data will be meta-analysed using a Bayesian approach (using WinBugs software) if 3 or more data points are found. If meta-analysis is not possible, data will be presented as individual values in adapted GRADE profile tables and plots of un-pooled sensitivity and specificity from RevMan software.
16.	Analysis of sub-groups	*Non-conditional stratification* Age: young (<18) vs adults (≥18) Follow up times: <6 months, 6-12 months, 1-5 years, >5 years *Conditional stratification* If heterogeneity is identified, where data is available, subgroup analysis will be carried out for the following subgroups: • Young subgroups: <2, 2-11, 11-18; Adults: 18-55, >55 • Learning disability versus no learning disability • Head injury vs no head injury • Type of epilepsy • gender
17.	Type and method of review	☐ Intervention ☐ Diagnostic ☒ Prognostic ☐ Qualitative ☐ Epidemiologic ☐ Service Delivery ☐ Other (please specify)
18.	Language	English

19.	Country	England		
20.	Anticipated or actual start date			
21.	Anticipated completion date			
22.	Stage of review at time of this submission	**Review stage**	**Started**	
		Preliminary searches	☐	
		Piloting of the study selection process	☐	
		Formal screening of search results against eligibility criteria	☐	
		Data extraction	☐	
		Risk of bias (quality) assessment	☐	
		Data analysis	☐	
23.	Named contact	**5a. Named contact** National Guideline Centre **5b Named contact e-mail** NGCEpilepsies@nice.org.uk **5e Organisational affiliation of the review** National Institute for Health and Care Excellence (NICE) and the National Guideline Centre		
24.	Review team members	National Guideline Centre:		
25.	Funding sources/sponsor	This systematic review is being completed by the National Guideline Centre which receives funding from NICE.		
26.	Conflicts of interest	All guideline committee members and anyone who has direct input into NICE guidelines (including the evidence review team and expert witnesses) must declare any potential conflicts of interest in line with		

		NICE's code of practice for declaring and dealing with conflicts of interest. Any relevant interests, or changes to interests, will also be declared publicly at the start of each guideline committee meeting. Before each meeting, any potential conflicts of interest will be considered by the guideline committee Chair and a senior member of the development team. Any decisions to exclude a person from all or part of a meeting will be documented. Any changes to a member's declaration of interests will be recorded in the minutes of the meeting. Declarations of interests will be published with the final guideline.
27.	Collaborators	Development of this systematic review will be overseen by an advisory committee who will use the review to inform the development of evidence-based recommendations in line with section 3 of Developing NICE guidelines: the manual. Members of the guideline committee are available on the NICE website: https://www.nice.org.uk/guidance/indevelopment/gid-ng10112.
28.	Other registration details	N/A
29.	Reference/URL for published protocol	
30.	Dissemination plans	NICE may use a range of different methods to raise awareness of the guideline. These include standard approaches such as: • notifying registered stakeholders of publication • publicising the guideline through NICE's newsletter and alerts • issuing a press release or briefing as appropriate, posting news articles on the NICE website, using social media channels, and publicising the guideline within NICE.
31.	Keywords	Epilepsies, risk factors, seizure
32.	Details of existing review of same topic by same authors	N/A
33.	Current review status	☒ Ongoing ☐ Completed but not published ☐ Completed and published ☐ Completed, published and being updated ☐ Discontinued
35.	Additional information	N/A
36.	Details of final publication	www.nice.org.uk

A.2 Health economic review protocol

Review question	All questions – health economic evidence
Objectives	To identify health economic studies relevant to any of the review questions.
Search criteria	• Populations, interventions and comparators must be as specified in the clinical review protocol above. • Studies must be of a relevant health economic study design (cost–utility analysis, cost-effectiveness analysis, cost–benefit analysis, cost–consequences analysis, comparative cost analysis). • Studies must not be a letter, editorial or commentary, or a review of health economic evaluations. (Recent reviews will be ordered although not reviewed. The bibliographies will be checked for relevant studies, which will then be ordered.) • Unpublished reports will not be considered unless submitted as part of a call for evidence. • Studies must be in English.
Search strategy	A health economic study search will be undertaken using population-specific terms and a health economic study filter.
Review strategy	Studies not meeting any of the search criteria above will be excluded. Studies published before 2004, abstract-only studies and studies from non-OECD countries or the USA will also be excluded. Studies published after 2004 that were included in the previous guideline(s) will be reassessed for inclusion and may be included or selectively excluded based on their relevance to the questions covered in this update and whether more applicable evidence is also identified. Each remaining study will be assessed for applicability and methodological limitations using the NICE economic evaluation checklist which can be found in appendix H of Developing NICE guidelines: the manual (2014).[8] **Inclusion and exclusion criteria** • If a study is rated as both 'Directly applicable' and with "Minor limitations" then it will be included in the guideline. A health economic evidence table will be completed, and it will be included in the health economic evidence profile. • If a study is rated as either 'Not applicable' or with "Very serious limitations" then it will usually be excluded from the guideline. If it is excluded, then a health economic evidence table will not be completed, and it will not be included in the health economic evidence profile. • If a study is rated as 'Partially applicable', with 'Potentially serious limitations' or both then there is discretion over whether it should be included. **Where there is discretion** The health economist will make a decision based on the relative applicability and quality of the available evidence for that question, in discussion with the guideline committee if required. The ultimate aim is to include health economic studies that are helpful for decision-making in the context of the guideline and the current NHS setting. If several studies are considered of sufficiently high applicability and methodological quality that they could all be included, then the health economist, in discussion with the committee if required, may decide to include only the most applicable studies and to selectively exclude the remaining studies. All studies excluded on the basis of applicability or methodological limitations will be listed with explanation in the excluded health economic studies appendix below. The health economist will be guided by the following hierarchies.

	Setting: • UK NHS (most applicable). • OECD countries with predominantly public health insurance systems (for example, France, Germany, Sweden). • OECD countries with predominantly private health insurance systems (for example, Switzerland). • Studies set in non-OECD countries or in the USA will be excluded before being assessed for applicability and methodological limitations. *Health economic study type:* • Cost–utility analysis (most applicable). • Other type of full economic evaluation (cost–benefit analysis, cost-effectiveness analysis, cost–consequences analysis). • Comparative cost analysis. • Non-comparative cost analyses including cost-of-illness studies will be excluded before being assessed for applicability and methodological limitations. *Year of analysis:* • The more recent the study, the more applicable it will be. • Studies published in 2004 or later (including any such studies included in the previous guideline(s)) but that depend on unit costs and resource data entirely or predominantly from before 2004 will be rated as 'Not applicable'. • Studies published before 2004 (including any such studies included in the previous guideline(s)) will be excluded before being assessed for applicability and methodological limitations. *Quality and relevance of effectiveness data used in the health economic analysis:* • The more closely the clinical effectiveness data used in the health economic analysis match with the outcomes of the studies included in the clinical review the more useful the analysis will be for decision-making in the guideline.

Appendix B Literature search strategies

This literature search strategy was used for the following reviews:

- What are the most accurate tools for predicting a further seizure, in people who have had a single seizure?
- What are the most accurate tools to predicting death, including SUDEP, in people with epilepsy?

The literature searches for this review are detailed below and complied with the methodology outlined in Developing NICE guidelines: the manual.[8]

For more information, please see the Methodology review published as part of the accompanying documents for this guideline.

B.1 Clinical search literature search strategy

Searches were constructed using the following approach:

- Population AND risk factor terms

Table 6: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline (OVID)	1946 – 13 August 2020	Exclusions
Embase (OVID)	1974 – 13 August 2020	Exclusions

Medline (Ovid) search terms

1.	exp epilepsy/
2.	seizures/
3.	exp status epilepticus/
4.	seizures, febrile/
5.	(dravet syndrome or epilep* or convuls* or continuous spike wave or slow sleep or landau kleffner syndrome or lennox gastaut syndrome or infant* spasm* or seizure* or west syndrome).ti,ab.
6.	or/1-5
7.	letter/
8.	editorial/
9.	news/
10.	exp historical article/
11.	Anecdotes as Topic/
12.	comment/
13.	case report/
14.	(letter or comment*).ti.
15.	or/7-14
16.	randomized controlled trial/ or random*.ti,ab.
17.	15 not 16
18.	animals/ not humans/
19.	exp Animals, Laboratory/
20.	exp Animal Experimentation/
21.	exp Models, Animal/
22.	exp Rodentia/

23.	(rat or rats or mouse or mice).ti.
24.	or/17-23
25.	6 not 24
26.	limit 25 to English language
27.	(risk* adj2 assess*).ti,ab.
28.	((score* or scoring) adj2 (tool* or system*)).ti,ab.
29.	((risk* or predict* or prognos*) adj4 (tool* or rule* or index* or indices or score* or scoring or scale* or model* or system* or algorithm* or stratif* or criteria or calculat*)).ti,ab.
30.	("ERA scale" or "ERA checklist" or "ERAC" or "Epilepsy risk awareness scale" or "SUDEP and seizure safety checklist" or "Epilepsy self-management scale" or "ESMS" or "Chalfont Seizure Severity Scale").ti,ab.
31.	((risk or predict*) and "EpSMon").ti,ab.
32.	or/27-31
33.	26 and 32

Embase (Ovid) search terms

1.	exp epilepsy/
2.	seizure/
3.	epileptic state/
4.	febrile convulsion/
5.	(dravet syndrome or epilep* or convuls* or continuous spike wave or slow sleep or landau kleffner syndrome or lennox gastaut syndrome or infant* spasm* or seizure* or west syndrome).ti,ab.
6.	or/1-5
7.	letter.pt. or letter/
8.	note.pt.
9.	editorial.pt.
10.	case report/ or case study/
11.	(letter or comment*).ti.
12.	or/7-11
13.	randomized controlled trial/ or random*.ti,ab.
14.	12 not 13
15.	animal/ not human/
16.	nonhuman/
17.	exp Animal Experiment/
18.	exp Experimental Animal/
19.	animal model/
20.	exp Rodent/
21.	(rat or rats or mouse or mice).ti.
22.	or/14-21
23.	6 not 22
24.	limit 23 to English language
25.	(risk* adj2 assess*).ti,ab.
26.	((score* or scoring) adj2 (tool* or system*)).ti,ab.
27.	((risk* or predict* or prognos*) adj4 (tool* or rule* or index* or indices or score* or scoring or scale* or model* or system* or algorithm* or stratif* or criteria or calculat*)).ti,ab.

28.	("ERA scale" or "ERA checklist" or "ERAC" or "Epilepsy risk awareness scale" or "SUDEP and seizure safety checklist" or "Epilepsy self-management scale" or "ESMS" or "Chalfont Seizure Severity Scale").ti,ab.
29.	((risk or predict*) and "EpSMon").ti,ab.
30.	or/25-29
31.	24 and 30

B.2 Health Economics literature search strategy

Health economic evidence was identified by conducting a broad search relating to an Epilepsies population in NHS Economic Evaluation Database (NHS EED – this ceased to be updated after March 2015) and the Health Technology Assessment database (HTA) with no date restrictions. NHS EED and HTA databases are hosted by the Centre for Research and Dissemination (CRD). Additional searches were run on Medline and Embase for health economics and quality of life studies.

Table 7: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline	Health Economics 1 January 2014 – 13 May 2021 Quality of Life 1946 – 13 May 2021	Health economics studies Quality of life studies Exclusions
Embase	Health Economics 1 January 2014 – 13 May 2021 Quality of Life 1974 – 13 May 2021	Health economics studies Quality of life studies Exclusions
Centre for Research and Dissemination (CRD)	HTA - Inception – 13 May 2021 NHSEED - Inception to 31 March 2015	None

Medline (Ovid) search terms

1.	exp epilepsy/
2.	seizures/
3.	exp status epilepticus/
4.	seizures, febrile/
5.	(dravet syndrome or epilep* or continuous spike wave or slow sleep or landau kleffner syndrome or lennox gastaut syndrome or infant* spasm* or seizure* or west syndrome).ti,ab.
6.	or/1-5
7.	letter/
8.	editorial/
9.	news/
10.	exp historical article/
11.	Anecdotes as Topic/
12.	comment/
13.	case report/
14.	(letter or comment*).ti.
15.	or/7-14

16.	randomized controlled trial/ or random*.ti,ab.
17.	15 not 16
18.	animals/ not humans/
19.	exp Animals, Laboratory/
20.	exp Animal Experimentation/
21.	exp Models, Animal/
22.	exp Rodentia/
23.	(rat or rats or mouse or mice).ti.
24.	or/17-23
25.	6 not 24
26.	limit 25 to English language
27.	Economics/
28.	Value of life/
29.	exp "Costs and Cost Analysis"/
30.	exp Economics, Hospital/
31.	exp Economics, Medical/
32.	Economics, Nursing/
33.	Economics, Pharmaceutical/
34.	exp "Fees and Charges"/
35.	exp Budgets/
36.	budget*.ti,ab.
37.	cost*.ti.
38.	(economic* or pharmaco?economic*).ti.
39.	(price* or pricing*).ti,ab.
40.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
41.	(financ* or fee or fees).ti,ab.
42.	(value adj2 (money or monetary)).ti,ab.
43.	or/27-42
44.	quality-adjusted life years/
45.	sickness impact profile/
46.	(quality adj2 (wellbeing or well being)).ti,ab.
47.	sickness impact profile.ti,ab.
48.	disability adjusted life.ti,ab.
49.	(qal* or qtime* or qwb* or daly*).ti,ab.
50.	(euroqol* or eq5d* or eq 5*).ti,ab.
51.	(health utility* or utility score* or disutilit* or utility value*).ti,ab.
52.	(hui or hui1 or hui2 or hui3).ti,ab.
53.	(health* year* equivalent* or hye or hyes).ti,ab.
54.	discrete choice*.ti,ab.
55.	rosser.ti,ab.
56.	(willingness to pay or time tradeoff or time trade off or tto or standard gamble*).ti,ab.
57.	(sf36* or sf 36* or short form 36* or shortform 36* or shortform36*).ti,ab.
58.	(sf20 or sf 20 or short form 20 or shortform 20 or shortform20).ti,ab.
59.	(sf12* or sf 12* or short form 12* or shortform 12* or shortform12*).ti,ab.
60.	(sf8* or sf 8* or short form 8* or shortform 8* or shortform8*).ti,ab.
61.	(sf6* or sf 6* or short form 6* or shortform 6* or shortform6*).ti,ab.

62.	or/44-61
63.	26 and (43 or 62)

Embase (Ovid) search terms

1.	exp *epilepsy/
2.	*landau kleffner syndrome/
3.	exp *seizure/
4.	"seizure, epilepsy and convulsion"/
5.	(dravet syndrome or epilep* or continuous spike wave or slow sleep or landau kleffner syndrome or lennox gastaut syndrome or infant* spasm* or seizure* or west syndrome).ti,ab.
6.	or/1-5
7.	letter.pt. or letter/
8.	note.pt.
9.	editorial.pt.
10.	case report/ or case study/
11.	(letter or comment*).ti.
12.	or/7-11
13.	randomized controlled trial/ or random*.ti,ab.
14.	12 not 13
15.	animal/ not human/
16.	nonhuman/
17.	exp Animal Experiment/
18.	exp Experimental Animal/
19.	animal model/
20.	exp Rodent/
21.	(rat or rats or mouse or mice).ti.
22.	or/15-21
23.	6 not 22
24.	limit 23 to English language
25.	health economics/
26.	exp economic evaluation/
27.	exp health care cost/
28.	exp fee/
29.	budget/
30.	funding/
31.	budget*.ti,ab.
32.	cost*.ti.
33.	(economic* or pharmaco?economic*).ti.
34.	(price* or pricing*).ti,ab.
35.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
36.	(financ* or fee or fees).ti,ab.
37.	(value adj2 (money or monetary)).ti,ab.
38.	or/25-37
39.	quality adjusted life year/
40.	sickness impact profile/

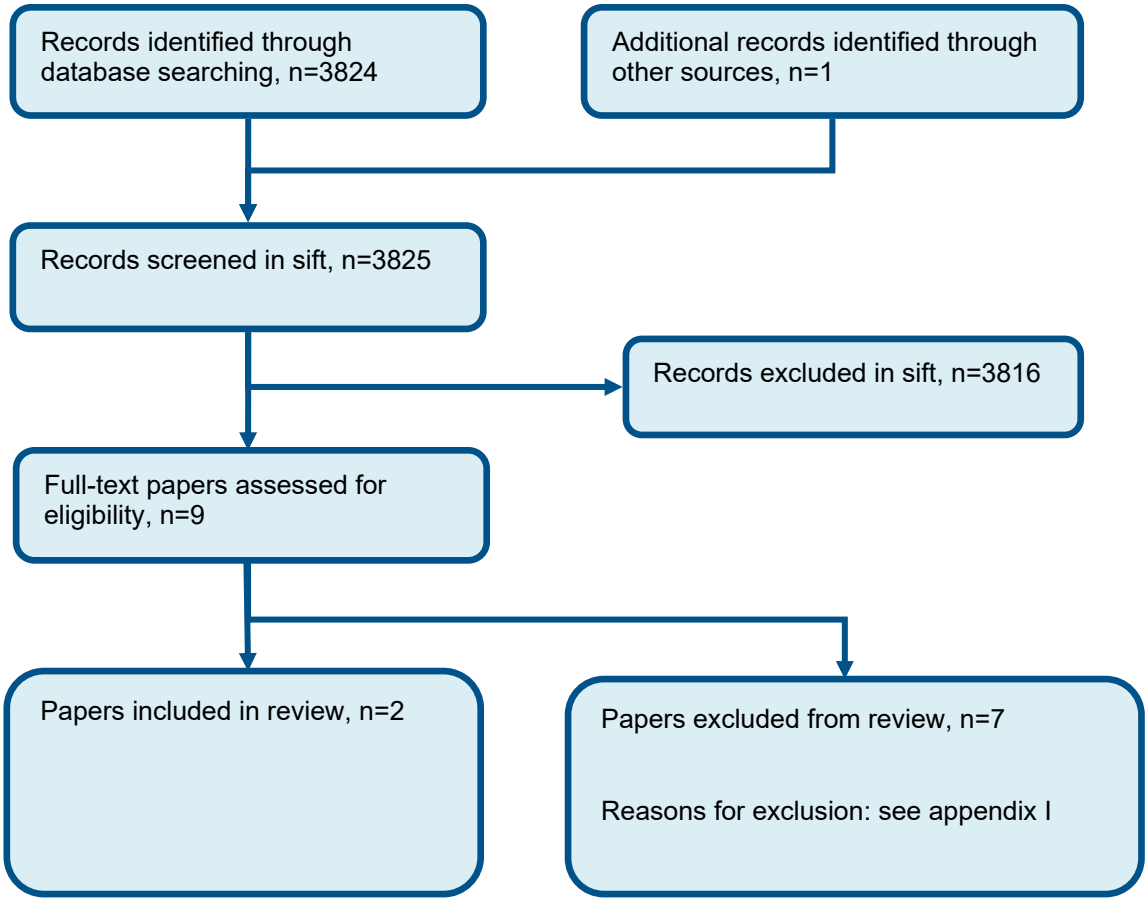
41.	(quality adj2 (wellbeing or well being)).ti,ab.
42.	sickness impact profile.ti,ab.
43.	disability adjusted life.ti,ab.
44.	(qal* or qtime* or qwb* or daly*).ti,ab.
45.	(euroqol* or eq5d* or eq 5*).ti,ab.
46.	(qol* or hql* or hqol* or h qol* or hrqol* or hr qol*).ti,ab.
47.	(health utility* or utility score* or disutilit* or utility value*).ti,ab.
48.	(hui or hui1 or hui2 or hui3).ti,ab.
49.	(health* year* equivalent* or hye or hyes).ti,ab.
50.	discrete choice*.ti,ab.
51.	rosser.ti,ab.
52.	(willingness to pay or time tradeoff or time trade off or tto or standard gamble*).ti,ab.
53.	(sf36* or sf 36* or short form 36* or shortform 36* or shortform36*).ti,ab.
54.	(sf20 or sf 20 or short form 20 or shortform 20 or shortform20).ti,ab.
55.	(sf12* or sf 12* or short form 12* or shortform 12* or shortform12*).ti,ab.
56.	(sf8* or sf 8* or short form 8* or shortform 8* or shortform8*).ti,ab.
57.	(sf6* or sf 6* or short form 6* or shortform 6* or shortform6*).ti,ab.
58.	or/39-57
59.	24 and (38 or 58)

NHS EED and HTA (CRD) search terms

#1.	MeSH DESCRIPTOR Epilepsy EXPLODE ALL TREES
#2.	MeSH DESCRIPTOR Seizures EXPLODE ALL TREES
#3.	MeSH DESCRIPTOR Status Epilepticus EXPLODE ALL TREES
#4.	MeSH DESCRIPTOR Seizures, Febrile EXPLODE ALL TREES
#5.	((dravet syndrome or epilep* or continuous spike wave or slow sleep or landau kleffner syndrome or lennox gastaut syndrome or infant* spasm* or seizure* or west syndrome))
#6.	#1 OR #2 OR #3 OR #4 OR #5

Appendix C Diagnostic evidence study selection

Figure 2: Flow chart of clinical study selection for the review of prediction of second seizure



Appendix D Predictive evidence

Reference	**Bonnett, 2014[3]**
Study type	External validation study
Study methodology	Data source: Group of unselected patients with newly diagnosed seizures from UK primary care system (NGPSE), adults referred to First Seizure Clinics of Royal Perth and Freemantle Hospitals (WA), and patients examined in 14 university clinics and hospitals in Italy with a first seizure (FIRST). Recruitment: External datasets used (retrospective)
Number of patients	n = 274 (NGPSE) + 847 (WA) + 305 (FIRST) = 1426
Patient characteristics	Age, median (IQR): 50.3 (31.8-68.8) [NGPSE]; 39 (26.0-56.0) [WA]; 28 (20.0-46.0) [FIRST] Gender (male to female ratio): 848:578 Ethnicity: unclear Setting: Primary and secondary care Country: UK, Italy, Australia Learning disability: unclear Head Injury: unclear Type of Epilepsy: unclear Inclusion criteria: NGPSE: People with newly diagnosed seizures, including children with febrile seizures (approximately 25%); 2/3 had definite or probable epilepsy; WA: People referred to a first seizure clinic in major teaching hospitals; FIRST: people having antiepileptic drug treatment after a first unprovoked tonic-clonic seizure Exclusion criteria: No specific criteria given for each site; those without outcome data
Target condition(s)	Epilepsy – second seizure

Reference	Bonnett, 2014[3]
Index test(s) and reference standard	Index predictive test Model developed by the MESS trial (Bonnett, 2010), which aims to predict recurrent seizure within 12 months of first seizure. The model includes variables for aetiology, epilepsy in a first degree relative, seizure while asleep, electroencephalogram (EEG) results, CT or MRI imaging and treatment policy. Risk groups were determined by the 16th, 505th and 84th centiles – thus low risk (0-16), moderate risk (16-50), high risk (50-84) and very high risk (84-100). Reference standard (and follow up) Second seizure within 12 months of first seizure
Results	**Discrimination: Harrels C statistic (95% CI)** **Results given separately for each different cohort** NGPSE: 0.60(0.55-0.65) WA: 0.59(0.56-0.62) FIRST: 0.65 (0.60-0.70). This was using multiple imputation to adjust for the fact that the sleep variable was missing. Other methods of imputation gave very similar results. Results are given for MESS, but this relates to the results in the developmental cohort [C=0.59 (0.56 – 0.63)] The paper did not provide a clear description of discrimination results. No data were given on the number of second seizures observed and there were no tables indicating the numbers with second seizure in each category of risk indicated by the MESS tool. **Calibration** The charts below have been pasted directly from the paper. The charts represent how well the predictions made by the MESS tool agree with the observed data on second seizures from each validation cohort. No quantitative data were provided by the paper.

Reference	Bonnett, 2014[3]

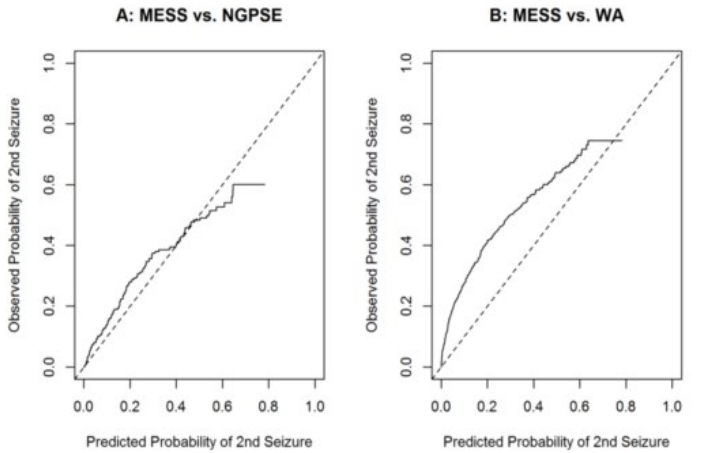


Figure 2. Calibration plots for MESS compared to NGPSE (A), WA (B), FIRST - variable matching (C) and FIRST - hot deck imputation (D).
doi:10.1371/journal.pone.0099063.g002

Reference	Bonnett, 2014[3]
Source of funding	Independent research funded by the National Institute for Health research. Declaration made of: no competing interests.
Limitations	Risk of bias: Serious; unclear if those adjudicating the tool score were aware of outcome. Indirectness: No serious indirectness
Comments	This was a poorly reported paper, although methodologically it appeared very sound. The MESS tool itself is inadequately described, and key information such as the number of second seizures observed was missing. There were also no tables indicating the numbers with second seizure in each category of risk indicated by the MESS tool. The lack of data on sensitivity and specificity would make it very difficult to recommend this tool – what are the relative contributions of sensitivity and specificity to the C statistic – is it dominated by sensitivity or otherwise? Sensitivity is probably the more important measure in this context.

Reference	Kim, 2016 [6]
Study type	Internal validation study
Study methodology	Data source: Ewha stroke registry Recruitment: External datasets used (retrospective)
Number of patients	n = 124
Patient characteristics	Age, median (IQR): 68 (57-75) Gender (male to female ratio): 69:55 Ethnicity: unclear Setting: Secondary care Country: South Korea Learning disability: No

Reference	Kim, 2016 [6]
	Head Injury: No Type of Epilepsy: 68 Generalised; 56 Partial; all post-stroke. Inclusion criteria: People experiencing post-stroke seizure after ischemic stroke (PSSi), who were on the Ewha Stroke Registry. Acute stroke, defined as acute neurological symptoms of presumed vascular origin lasting >24hrs and confirmed by radiologic findings. Seizure diagnosed clinically and distinguished as partial or generalised (ILAE). Exclusion criteria: TIA instead of acute stroke; primary haemorrhagic stroke; previous diagnosis of epilepsy; probable epileptogenic lesions; non convulsive EEG seizure; LOC or confusion only.
Target condition(s)	Epilepsy – second seizure
Index test(s) and reference standard	Index predictive test: Three tools were evaluated: 1. The study prediction tool comprised 10 predictors: seizure onset under 65yrs, male gender, AF, lesion size, cortex involvement, haemorrhagic transformation, functional disability after stroke, status epilepticus, stroke lesion EEG findings, partial seizure. 7 versions of these were tested (1-1, 1-2, 2-1, 2-2, 3-1, 3-2, 4; these are defined in the results section below) 2. Post-Stroke Epilepsy Risk Scale (PoSERS) tool (reference 32 - Strzelczyk): supratentorial, cortical involvement, haemorrhagic transformation, modified Rankin score ≥3 3. MESS tool: modified Rankin score ≥1, abnormal EEG Reference standard (and follow up) Unprovoked second seizure separated from first one by >24hrs. Mean follow up period to end point of study was 29.9 months post first seizure.
Results	54 second seizures reported within follow up of 29.9 months. Data in the study were stratified by whether the initial first seizure was 'early PSSi' (≤7 days between stroke and first seizure) or 'late PSSi' (>7 days between stroke and first seizure).

Reference	Kim, 2016 [6]
	Discrimination Early PSSi Score 1-1 (Y,M,A,B,C,H,D,E,F,P): AUC 0.653(0.502-0.784); sen 58.8%, spec 67.7% (6 = threshold) Score 1-2 (Y,M,A,B,C,H,mRS,E,F,P): AUC 0.650(0.499-0.782); sen 52.9%, spec 80.7% (9 = threshold) Score 2-1 (B,C,H,D,E,F,P): AUC 0.621(0.470-0.757); sen 64.7%, spec 71.0% (5 = threshold) Score 2-2 (B,C,H,mRS,E,F,P): AUC 0.622(0.471-0.758); sen 52.9%, spec 80.7% (8 = threshold) Score 3-1 (M, A, C, D, P): AUC 0.735(0.588-0.852); sen 70.6%, spec 71.0% (3 = threshold) Score 3-2 (M, A, C, mRS, P): AUC 0.676(0.525-0.803); sen 47.1%, spec 80.7% (6 = threshold) MESS (mRS≥1, abnormal EEG): AUC 0.509(0.361-0.657); sen 70.6%, spec 35.5% (1 = threshold) PoSERS (supratentorial, C,H,mRS≥3): AUC 0.576(0.425-0.717); sen 58.8%, spec 58.1% (3 = threshold) Late PSSi Score 1-1 (Y,M,A,B,C,H,D,E,F,P): AUC 0.566(0.447-0.679); sen 46.0%, spec 61.5% (7 = threshold) Score 1-2 (Y,M,A,B,C,H,mRS,E,F,P): AUC 0.558(0.439-0.671); sen 40.5%, spec 69.2% (9 = threshold) Score 2-1 (B,C,H,D,E,F,P): AUC 0.534(0.416-0.649); sen 43.2%, spec 59.0% (5 = threshold) Score 2-2 (B,C,H,mRS,E,F,P): AUC 0.526(0.408-0.642); sen 54.1%, spec 59.0% (7 = threshold) Score 4 (Y,M,B, P): AUC 0.734(0.620-0.829); sen 62.2%, spec 81.1% (3 = threshold) MESS (mRS≥1, abnormal EEG): AUC 0.594(0.475-0.705); sen 32.4%, spec 87.2% (1 threshold) PoSERS (supratentorial, C,H,mRS≥3): AUC 0.532(0.414-0.647); sen 21.6%, spec 89.7% (4 threshold) **Y=younger age <65[1 pt]; M=male [1 pt]; A=AF [1 pt]; B=bigger lesion size [small=0pt, moderate=1pt, large=2pt]; C=cortical involvement [1]; H=haemorrhagic transformation [1pt]; D=functional**

Reference	**Kim, 2016 [6]**
	disability[mild=opt, moderate=1pt, severe=2pt]; E=status epilepticus[1pt]; F=relevant focal EEG finding[1pt]; P=partial seizure [1pt].
Source of funding	Korea Health technology R&D Project through Korea Health Industry Development Institute, funded by Ministry of Health and Welfare Basic science Research Program through National Research Foundation of Korea, funded by the Ministry of Education, Science and technology
Limitations	Risk of bias: Serious; unclear if those adjudicating the tool score were aware of outcome. Indirectness: No serious indirectness
Comments	Post-stroke seizure

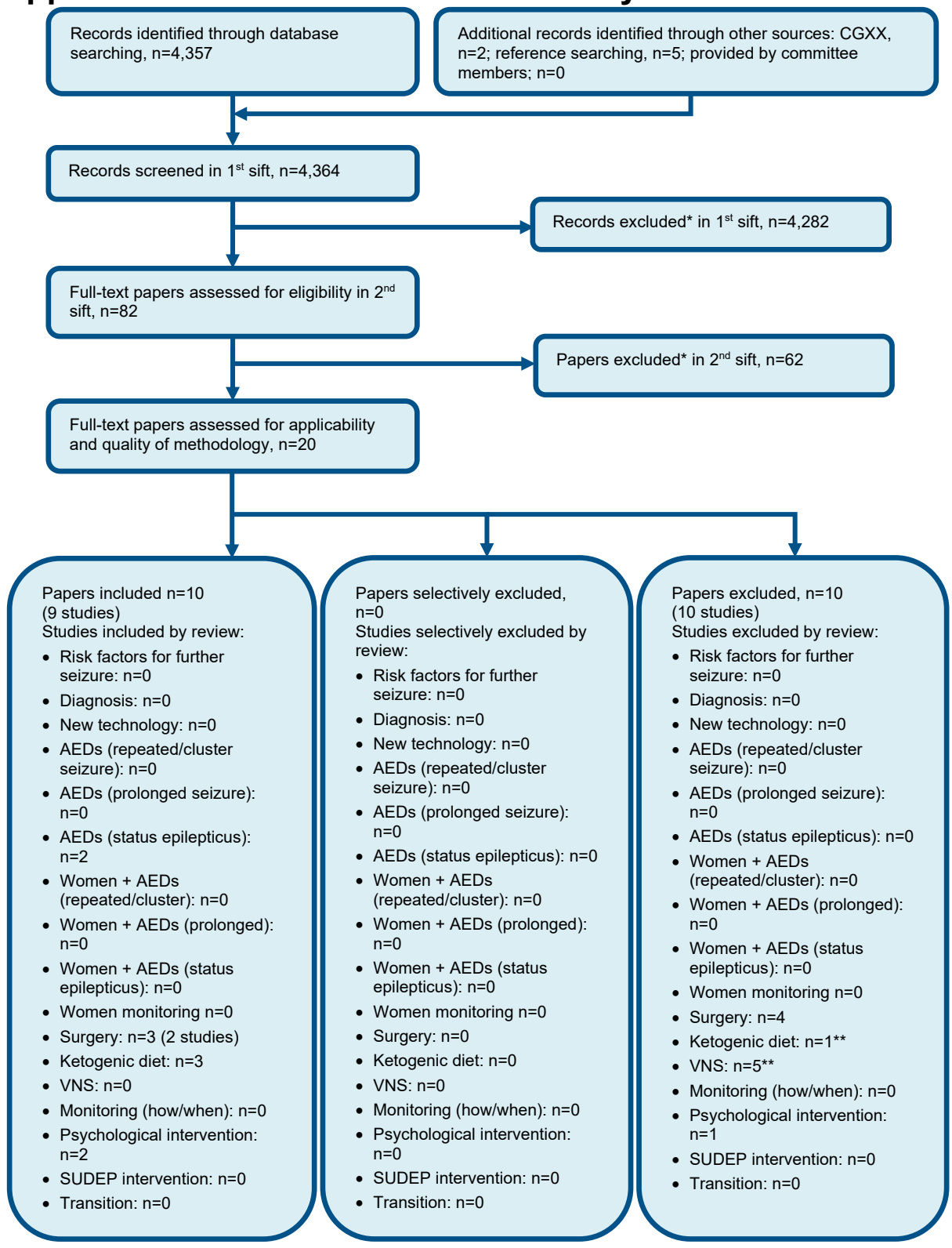
Appendix E Risk of bias (PROBAST)

Study	Appropriate data sources?	Appropriate inclusions and exclusions?	Similar health across participants?	Predictors defined/assessed same for all?	Predictor assessments made without knowledge of outcome data?	Predictors all available at time model meant to be used?	All relevant predictors analysed?	Pre-specified outcome used?	Predictors excluded from outcome definition?	Outcome defined in same way for all?	Outcome determined without knowledge of predictor information?	Reasonable number of outcome events? (100)	Time interval between baseline and outcome appropriate? (≥1 years)	All enrolled included in analysis?	Missing data handled appropriately?	Non-binary predictors handled appropriately?	Complexities in data accounted for?	Relevant performance measures?	Model recalibrated or likely that calibration not needed?	Overall rating
Bonnett, 2014[3]	Y	Y	Y	Y	U	N	N	Y	Y	Y	Y	U	Y	Y	Y	Y	Y	Y	Y	Serious risk of bias
Kim, 2016[6]	Y	Y	Y	Y	U	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Serious risk of bias

Appendix F Forest plots

Not applicable

Appendix G Economic evidence study selection



* Non-relevant population, intervention, comparison, design or setting; non-English language
**Please note that 1 article related to two questions. For this reason, the numbers listed for each review may not total the number of full text articles assessed for applicability and quality of methodology.

Appendix H Economic evidence tables

None.

Appendix I Health economic model

No original health economic modelling was undertaken for this review question.

Appendix J Excluded studies

J.1 Clinical studies

Table 8: Studies excluded from the clinical review

Study	Reason for exclusion
Kim, 2006[7]	Prediction tool tested on a sample that were not predominantly with one previous seizure – most had experienced multiple previous seizures.
Choquet, 2008 [5]	Prediction tool tested on a sample that were not predominantly with one previous seizure – most had experienced multiple previous seizures.
Anonymous, 2014[1]	No data - correction of graph in Bonnett 2014
Bonnett, 2010[4]	no useable outcomes
Berg, 1991[2]	review - references checked
Strzelczyk, 2010[9]	Not for prediction after first seizure - purely for prediction of seizure post stroke
Tews, 2015 [10]	Did not evaluate a prediction tool. This paper is included in the diagnostic accuracy review.

(a) <Insert Note here>

J.2 Health Economic studies

Published health economic studies that met the inclusion criteria (relevant population, comparators, economic study design, published 2004 or later and not from non-OECD country or USA) but that were excluded following appraisal of applicability and methodological quality are listed below. See the health economic protocol for more details.

Table 9: Studies excluded from the health economic review

Reference	Reason for exclusion
None.	

Appendix K Research recommendations

K.1.1 A prediction tool developmental study: this will be a large-scale cohort study, including only people with a first seizure at baseline to better determine the likelihood of an individual experiencing a second seizure. A suitable title for this research study could be: '*Development of a risk prediction tool to detect second seizure, in a cohort of people with a single seizure, using logistic regression modelling'*.

K.1.2 An external validation study: this will involve testing the risk prediction tool(s) developed by the previous study. A suitable title for this research study could be: '*External validation of a risk prediction tool to detect the probability of a second seizure in people with a single seizure at baseline'*.

Why this is important

One in twenty people will experience a seizure at some stage in their lifetime. In some people, there is evidence of an enduring predisposition to further events, for example, if a person has had a stroke or a major head injury. In many people, it is not known whether they will have a second seizure. Predicting who will have a second seizure could be important in mitigating an individual's risk and, at a population level, can help influence regulations in the workplace and for driving.

Rationale for research recommendation

Importance to 'patients' or the population	One of the most difficult things about seizures is their unpredictability. Given how common isolated seizures are, it would be of broad benefit to better know who may experience a second attack and within what timeframe. While this may not immediately lead to prescribing anti-seizure medications, armed with better models of who is at higher risk of further events, there can be a clear discussion about how to avoid provoking factors for seizures and how to minimise the risk were a second seizure to occur.
Relevance to NICE guidance	Prediction of a second seizure has been considered in this guideline, and we were unable to find evidence to support tools to predict a second seizure. To develop such tools, a well-curated database of people who have experienced a single seizure is essential.
Relevance to the NHS	The outcome would shape discussions with people who have had an isolated seizure and could influence medical investigations and the order in which investigations are performed. The data acquired would further help in treatment decision making and could have a broader impact, for example, on driving regulations.
National priorities	Moderate to high

Current evidence base	Very minimal, low to very low-quality data are currently available
Equality considerations	The data are likely to be equally applicable across sex and ethnicities. There should be specific work on people with intellectual disabilities and the elderly. These are groups in which the prevalence of epilepsy is higher, and the risks from seizures can be greater.

Modified PICO table

Population	People who have experienced a first unprovoked seizure
Intervention	Application of second seizure risk prediction tool, the tool being based on a large dataset of well phenotyped individuals who have had a first seizure.
Comparator	Standard care
Outcome	Accurate prediction of a second seizure Subsequent work would determine how to better prevent second seizures
Study design	Initial longitudinal study and subsequent clinical trial/ validation study of the prediction tool.
Timeframe	Long term
Additional information	None