BMJ Paediatrics Open

Prediction of recurrent febrile seizures risk during the same febrile illness in children at a single tertiary centre in Turkiye

Gül Yücel ,[1,2] Ahmet Kadir Arslan,[3] Bilge Özgör,[2] Serdal Güngör[2]

To cite: Yücel G, Arslan AK, Özgör B, *et al*. Prediction of recurrent febrile seizures risk during the same febrile illness in children at a single tertiary centre in Turkiye. *BMJ Paediatrics Open* 2025;**9**:e002908. doi:10.1136/bmjpo-2024-002908

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bmjpo-2024-002908).

Received 11 July 2024
Accepted 30 March 2025

[1]Inonu University, Malatya, Turkey
[2]Pediatric Neurology, İnönü Üniversitesi Tıp Fakültesi, Malatya, Turkey
[3]Department of Biostatistics and Medical Informatics, İnönü Üniversitesi Tıp Fakültesi, Malatya, Turkey

Correspondence to
Dr Gül Yücel; gul.yucel@inonu.edu.tr

ABSTRACT

Background This study aimed to develop a risk prediction model based on association rule mining to predict recurrent febrile seizures (RFS).

Methods This is a retrospective observational study that examined the medical records of 105 children who were followed up with febrile seizure (FS) in a tertiary paediatric emergency department between October 2022 and December 2023. Children were divided into RFS and simple FS groups. RFS was defined as seizures occurring more than once within 24 hours of the first FS in the same febrile illness. Risk factors associated with RFS were determined by univariate and multivariate analyses. χ^2, Mann-Whitney U, receiver operating characteristics (ROC), multiple logistic regression and Classification Based on Association Rules Algorithm (CBA) analyses were applied to the dataset to obtain high-level outputs.

Results RFS was detected in 32 out of 105 cases with FS (30.5%). Potential risk factors contributing to the development of RFS were seizure duration, number of recurrent seizures, family history, body temperature, time from fever onset to seizure, time from seizure onset to arrival at the emergency department, hyponatraemia, osmotic pressure and low haemoglobin level. The CBA algorithm obtained a total of 11 classification rules for the two patient groups. Additionally, the cut-off values obtained from CBA and ROC analysis showed satisfactory consistency. The CBA model achieved 97% overall accuracy classification performance.

Conclusion The developed CBA model shows good predictive ability for RFS. The relevant model can be used as a risk estimation tool to identify children at risk of developing RFS.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Recurrent febrile seizures (RFS) occur in some children.
⇒ Recognising risk factors associated with RFS can be useful in estimating risk probabilities and identifying high-risk cases.
⇒ Studies on developing risk prediction models are not at the desired level, or different modelling approaches are needed.

WHAT THIS STUDY ADDS

⇒ In this study, we proposed an approach to predict the risk of RFS by generating well-defined rules based on potential risk factors.
⇒ As can be seen from the overall accuracy, sensitivity and specificity values, the Classification Based on Association Rules Algorithm (CBA) model shows good predictive ability. Additionally, the interpretable rules obtained from this model have clinical application value.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The developed CBA model shows good predictive ability for RFS.
⇒ This model can be used as a risk prediction tool to identify children at risk of febrile seizure recurrence.

INTRODUCTION

Febrile seizures (FS) are the most common seizure type in childhood and usually occur in children between the ages of 6 months and 5 years.[1] The prevalence of FS in the USA and Western Europe is 2%–5%, but in Japan and Korea it is reported to be 6%–11%.[1–4] Seizures occurring more than once within 24 hours of the first FS during the same febrile illness are defined as recurrent FS (RFS).[5] This subgroup accounts for approximately 14%–24% of patients with FS.[5–12]

There are three concerns regarding RFS in the emergency department (ED): First, whether the fever is due to a serious illness such as bacterial meningitis or acute encephalitis; second, whether the seizure is the onset of epilepsy and will recur in the future; and third, the possibility of a new seizure occurring during the same illness process.[11 12] Following publication of guidelines by the American Academy of Pediatrics, children with FS have observed significant reductions in diagnostic testing, hospitalisations and healthcare costs without delay in the diagnosis of bacterial meningitis.[13 14] Additionally, since the introduction of *Haemophilus influenzae* type b and pneumococcal vaccines, the proportion of

patients diagnosed with bacterial meningitis in complex FS (CFS) (0.3%–0.7%) and the frequency of seizure recurrence during hospitalisation (3.8%) have decreased significantly.[15–17] In Japan, FS recurrence occurred in 16% of children with FS within 24 hours. The majority of seizure recurrences (82%) occurred within 8 hours of the initial seizure. This was found to be related to diazepam administration and a family history of FSs.[17] Short-term FSs are not considered to increase the risk of neurological or cognitive impairment, but there is a small increased risk of developing afebrile seizures.[18] The incidence of future epilepsy in children with CFS has been reported to be approximately 6%–8%. The risk of epilepsy is higher in focal features (6.3%) and prolonged seizures (29.4%). However, the risk of future epilepsy is lowest in seizures that recur within 24 hours (3.6%).[19]

The importance of accurate diagnosis in medical science necessitates the effective use of the rich wealth of information in databases. In this context, data mining is a powerful tool used to reveal hidden relationships and patterns by deeply analysing medical data. In particular, association rule mining is a frequently preferred method for discovering connections between items in large datasets.[20 21] Association rules specify conditions that are frequently observed together in a given dataset. Inferred rules explain the existence of certain features based on other features.[22 23] The Apriori algorithm is a frequently used method in the field of association rule mining and detects frequent item clusters in the dataset by determining certain support and confidence thresholds. This algorithm contributes to a deeper understanding of medical data by ensuring that the results obtained are interpretable and understandable.[24] In the field of paediatrics, the Classification Based on Association Rules Algorithm (CBA) has been used to develop various prediction models, helping to achieve important goals such as early diagnosis of diseases and optimisation of treatment processes.[25–27]

This study aimed to develop a risk prediction model for FS recurrence in children by generating well-defined rules based on association rule mining. The rules obtained from the CBA model may provide clinicians with an easily interpretable/applicable decision support system to estimate the risk of RFS.

MATERIALS AND METHODS

Study design and participants

Medical records of patients aged 6–60 months hospitalised with FS at a tertiary university hospital between October 2022 and December 2023 were retrospectively examined.

Inclusion criteria: (1) children aged 6 months to 5 years who met the criteria defined for FS and (2) hospitalisation of at least 24 hours

Exclusion criteria: (1) children with previous FSs and unprovoked seizures, (2) patients diagnosed with epilepsy, (3) having seizures due to central nervous system infection, gastroenteritis and congenital metabolic diseases, (4) patients with previous diagnoses such as genetic abnormality, head trauma, brain tumour, hydrocephalus, intracranial haemorrhage, developmental delay and cerebral palsy, (5) patients with incomplete medical records (demographic, clinical and laboratory) and (6) children referred from the primary care centre following a second seizure episode or children of parents who were brought to the ED following the second seizure episode were not included in the study.

Data on seizure types, duration of the first seizure, and the time from seizure to arrival at the ED were determined from interviews with parents and/or phone camera recordings.

Patient and public involvement

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

Definitions

Simple FS (SFS), in accordance with the definitions established by the International League Against Epilepsy, was defined as a brief generalised seizure lasting<15 min, not recurring within 24 hours, occurring during a febrile attack not caused by acute nervous system disease in a child between the ages of 6 months and 5 years without neurological deficits.[28]

CFS was defined as a focal or generalised and prolonged seizure lasting>15 min, recurring more than once within 24 hours and/or associated with post-seizure neurological abnormalities, more often post-seizure paralysis.[28]

RFS was defined as seizures occurring more than once within 24 hours during the same febrile illness.[5 11]

Data collection

Patient data were obtained from medical records and recorded on a data collection form. (1) Demographic information (patient age (months), gender, family history of FS, neurodevelopmental assessment), (2) characteristics of FS (body temperature before or immediately after the FS attack, FS duration, time from fever onset to FS, time from seizure onset to reaching ED, FS type, number of FS recurrences in 24 hours), (3) laboratory parameters at admission (white blood cells, neutrophils, lymphocytes, platelets, neutrophil–lymphocyte ratio (NLR), platelet–lymphocyte ratio, mean platelet volume–platelet count ratio, systemic immune inflammation index, haemoglobin, C-reactive protein, serum glucose, serum sodium, calculated osmotic pressure, magnesium) were collected. Osmotic pressure was calculated using the formula: osmotic pressure (mOsm/kg H2O) = sodium (mEq/L) × 2 + glucose (mg/dL)/18+-blood urea nitrogen (mg/dL)/2.8.[29]

Statistical analysis

In this study, to obtain outputs with a high level of evidence, χ^2, Mann-Whitney U, receiver operating characteristics (ROC), multiple logistic regression and CBA

Table 1 Cross table including distribution of some qualitative independent features in terms of febrile seizure types

Features*	Categories	Types of febrile seizure		Total (n=105)	χ^2 statistics	P value
		Simple febrile seizure (n=73)	Recurrent febrile seizure (n=32)			
Gender	Girl	32[a] (43.80)	17[a] (53.10)	49 (46.70)	0.443	0.506†
	Boy	41[a] (56.20)	15[a] (46.90)	56 (53.30)		
Duration of the first seizure (min)	≤1	55[a] (75.34)	0[b] (0.00)	55 (52.38)	63.67	**<0.001‡**
	>1 to <3	13[a] (17.81)	9[a] (28.13)	22 (20.95)		
	≥3 to <5	5[a] (6.85)	11[b] (34.38)	16 (15.24)		
	≥5 to <10	0[a] (0.00)	7[b] (21.88)	7 (6.67)		
	≥10 to 15	0[a] (0.00)	5[b] (15.60)	5 (4.80)		
Number of seizure recurrences within 24 hours	1	73[a] (100.00)	0[b] (0.00)	73 (69.52)	105	**<0.001‡**
	2	0[a] (0.00)	23[b] (71.88)	23 (21.90)		
	3	0[a] (0.00)	7[b] (21.88)	7 (6.67)		
	4	0[a] (0.00)	2[b] (6.25)	2 (1.90)		
Family history of febrile seizure in first-degree relatives	No	57[a] (78.08)	14[b] (43.75)	71 (67.62)	14.2	**0.001‡**
	Febrile seizure	16[a] (21.92)	16[b] (50.00)	32 (30.48)		
	Epilepsy	0[a] (0.00)	2[b] (6.25)	2 (1.90)		
Total		73 (100.00)	32 (100.00)	105 (100.00)		

P values which are≤0.05 are reported in bold.
a,b 'Each superscript indicates a subset of febrile seizure categories whose column proportions are not significantly different'.
*The features are summarised as frequency (column per cent).
†Pearson χ^2 with Yates' continuity correction.
‡Pearson χ^2 (exact).

analyses were applied to the dataset, respectively. The statistical analysis stage is presented under four subtitles.

A priori power analysis

A priori power analysis for detecting minimum sample size was performed based on the findings from the paper. In this context, the minimum sample size required to detect a significant difference using ROC analysis should be at least 52 in the control group and 11 in the case group (63 in total), considering type I error (alpha) of 0.05, power (1-beta) of 0.8, allocation ratio with 4.66 and difference of area under the curve (AUC) values (between observed AUC value and null hypothesis AUC value (0.50) 0.2672). The relevant calculation was made using the WSSPAS: Web-Based Sample Size and Power Analysis Software tool developed by Inonu University Faculty of Medicine, Department of Biostatistics and Medical Informatics.[30]

Univariate statistical analysis

The features used in this research were expressed with mean, SD, median, minimum and maximum values, frequency and percentage according to measurement levels. The evaluation of normality for quantitative features was achieved by the Shapiro-Wilk test. For quantitative features, to examine significant differences between FS groups, the Mann-Whitney U test was performed. For qualitative features, Pearson's χ^2 and Yates' χ^2 with continuity correction tests were used where appropriate. Comparisons of column proportions of FS types in terms of each category of row variables were performed with the Two-Proportions Z-Test. A series of ROC analyses were performed to select optimal cut-off values based on the balance between sensitivity and specificity metrics. Youden index was used to specify optimal cut-off values. Boxplots were generated to demonstrate the distribution of the relevant quantitative features. p≤0.05 was accepted as the statistical significance level.

Multivariate statistical analysis

To model the relationship between prognostic features and FS groups, multiple logistic regression analysis (MLRA) was performed. The Hosmer-Lemeshow test was employed to evaluate goodness of data-model fit. To select the best subset of features, the forward selection technique was used where feature removal was based on Wald statistics. The interpretations of the statistical relationship between FS types and relevant features were achieved by OR statistics.

In addition to multiple logistic regression, the CBA was used to provide an additional explainable perspective on prognostic features used to classify FS types.[31] CBA uses the knowledge embedded in association rules to enhance the classification process. Association rule mining is a technique in data mining used to discover interesting relationships or associations among a set of items in datasets. To mine classification rules, the Apriori

Table 2 The distributions of quantitative independent features in terms of febrile seizure types

Features*	Types of febrile seizure		Total (n=105)	P value†
	Simple febrile seizure (n=73)	Recurrent febrile seizure (n=32)		
Age (months)	23.78±13.45 | 21 (6–56)	27.25±12.86 | 26.5 (9–56)	24.84±13.31 | 24 (6–56)	0.131
Body temperature (°C)	39.3±0.76 | 39 (38–41)	38.47±0.63 | 38.5 (37.3–40)	39.05±0.82 | 38.9 (37.3–41)	**<0.001**
Time from onset of fever to seizure (hours)	4.38±1.17 | 4 (3–8)	12.91±5.51 | 12 (4–22)	6.98±5.05 | 5 (3–22)	**<0.001**
Time from onset of seizure to arrival at the ED (min)	33.63±4.03 | 34 (25–41)	36.16±5.22 | 37 (14–42)	34.4±4.56 | 35 (14–42)	**0.001**
Calculated osmotic pressure (mOsm/kg water)	287.68±5.06 | 286.4 (277.81–298)	281.05±5.97 | 280.39 (273.42–304.18)	285.66±6.15 | 285.2 (273.42–304.18)	**<0.001**
Systemic immune-inflammation index	728.4±754.63 | 440.73 (21.88–3682.68)	730.71±778.75 | 536.59 (117.48–4563.75)	729.1±758.3 | 473.99 (21.88–4563.75)	0.473
Mean platelet volume to platelet count ratio	9.55±0.88 | 9.6 (7.5–11.6)	9.54±0.92 | 9.2 (8.2–11.3)	9.55±0.89 | 9.5 (7.5–11.6)	0.725
Neutrophil to lymphocyte ratio	2.24±1.96 | 1.51 (0.24–12.32)	2.22±1.57 | 1.81 (0.23–7.8)	2.23±1.84 | 1.73 (0.23–12.32)	0.549
Platelet to lymphocyte ratio	89.74±47.92 | 79.58 (9.39–249.17)	100.82±72.82 | 84.34 (39.71–375)	93.11±56.54 | 81.09 (9.39–375)	0.749
C-reactive protein (mg/dL)	17.15±25.91 | 6.2 (3.02–150)	21.77±26.17 | 8.32 (3.1–104)	18.56±25.95 | 7.03 (3.02–150)	0.346
White blood cells ($\times 10^9$ /L)	12.98±6.68 | 11.28 (4.8–32.61)	12.21±4.24 | 11.59 (5.84–23.22)	12.74±6.03 | 11.28 (4.8–32.61)	0.821
Neutrophils (%)	7.5±5.57 | 5.43 (1.06–23.98)	7.11±3.51 | 6.23 (2.31–17.19)	7.38±5.02 | 5.66 (1.06–23.98)	0.446
Lymphocytes (%)	4.12±2.06 | 3.55 (1.1–9.81)	3.95±1.78 | 3.73 (0.75–9.91)	4.07±1.97 | 3.57 (0.75–9.91)	0.944
Platelets ($\times 10^9$ /L)	303.95±109.1 | 299 (43–625)	324.59±108.13 | 317.5 (147–585)	310.24±108.7 | 306 (43–625)	0.495
Red blood cell distribution width (%)	14.25±1.76 | 13.7 (12.1–21)	14.71±2.78 | 14.3 (4.8–21.2)	14.39±2.12 | 13.9 (4.8–21.2)	0.097
Haemoglobin (g/L)	118.8±11.5 | 117 (89–158)	112.6±12.4 | 112.5 (84–148)	116.9±12.1 | 116 (84–158)	**0.029**
Glucose (mg/dL)	113.41±31.52 | 106 (67–195)	110.09±36.05 | 99 (37–248)	112.4±32.82 | 104 (37–248)	0.519
Blood urea nitrogen (mg/dL)	22.74±7.77 | 21 (3–45)	24.84±13.79 | 21.5 (12–87)	23.38±9.97 | 21 (3–87)	0.983
Creatinine (mg/dL)	0.46±0.07 | 0.45 (0.32–0.7)	0.46±0.07 | 0.46 (0.34–0.75)	0.46±0.07 | 0.45 (0.32–0.75)	0.878
Aspartate aminotransferase (U/L)	40.41±10.47 | 38 (21–82)	39.88±9.77 | 37.5 (19–75)	40.25±10.22 | 38 (19–82)	0.928
Alanine aminotransferase (U/L)	18.26±11.48 | 15 (8–99)	16±4.46 | 15 (10–32)	17.57±9.91 | 15 (8–99)	0.807
Sodium (mEq/L)	136.63±1.51 | 137 (134–141)	133.03±1.2 | 133 (131–135)	135.53±2.19 | 136 (131–141)	**<0.001**
Potassium (mEq/L)	4.24±0.57 | 4.2 (3.1–5.8)	4.4±0.47 | 4.35 (3.4–5.9)	4.29±0.54 | 4.3 (3.1–5.9)	0.176
Calcium (mg/dL)	9.62±0.55 | 9.6 (8.1–10.8)	9.46±0.42 | 9.5 (8.6–10.4)	9.57±0.52 | 9.5 (8.1–10.8)	0.119
Magnesium (mg/dL)	2.07±0.19 | 2 (1.7–2.5)	2.1±0.23 | 2.1 (1.8–2.9)	2.08±0.2 | 2.1 (1.7–2.9)	0.637

P values which are≤0.05 are reported in bold.
*The features are summarised as mean±SD and median (min–max).
†Mann-Whitney U test.
ED, emergency department.

algorithm was implemented. In the CBA model, in order to constitute rule-based patterns based on quantitative features for outcome feature, the relevant features must be discretised.[32] For this purpose, Class-Attribute Contingency Coefficient (CACC) was used as the discretisation technique.[33]

To evaluate the interestingness of the classification rules, support and confidence metrics were used. Support measures the proportion of instances in the dataset that contain the items mentioned in the rule. Confidence measures the accuracy of the rule. It is the conditional probability of the consequent given the antecedent. The formulas of the related metrics given below:

$$Classification\ rule: A \rightarrow B$$

$$Support = \frac{Frequency(A,B)}{N}$$

$$Confidence = \frac{Frequency(A,B)}{Frequency(A)}$$

Accuracy, sensitivity and specificity, metrics were used to evaluate the performance of the CBA model in correctly classifying examples using learnt rules. Detailed information on how to calculate these metrics can be found in the paper.[34]

Statistical tools/softwares used

All statistical analyses of this research were performed using R (V.4.1.2) software with RStudio (V.2023.06.0) integrated development environment. Apart from basic analyses, pROC, arulesCBA, arulesViz, visNetwork, ggplot2 and ggstatsplot for featured analyses and graphs.

RESULTS

The results obtained by examining whether there was a statistically significant difference between the febrile groups in terms of the qualitative and quantitative features used in the study are presented in tables 1 and 2, respectively. According to table 1, there was a statistically significant difference between the FS groups in terms of 'Duration of seizure (min)', 'Number of seizure recurrent within 24 hours' and 'Family history of febrile seizure in first degree relatives' features. According to table 2, there was a statistically significant difference between the FS groups in terms of body temperature (°C), time from onset of fever to seizure (hours), time from onset of seizure to arrival at the ED (min), calculated osmotic pressure (mOsm/kg water), haemoglobin (g/L) and sodium (mEq/L). The features presented in table 2, which are statistically significant according to the Mann-Whitney U test, the boxplot graphs are given as merged in online supplemental figure 1.

The detailed outputs of ROC analysis with respect to the features which are statistically significant according to the Mann-Whitney U test are given in table 3. Additionally, area under the ROC curve graphs of related quantitative features are given as merged in figure 1.

The outputs of the MLRA applied to model the relationship between the features that were significant according to the Mann-Whitney U test and febrile seizure groups in order to obtain explainable outputs in terms of ORs are presented in table 4. According to table 4, Hosmer-Lemeshow results show that the model-data fit is significant (p=0.996). Since the forward selection technique was applied in the relevant MLRA model, only two features (body temperature (°C) and sodium (mEq/L)) with significant coefficients were observed. ORs for these two features were calculated as 0.057 and 0.043, respectively.

A total of 11 classification rules were obtained after the application of the CBA algorithm are given in table 5. While six of the antecedent rules predicted the type of febrile seizure as 'Simple', the remaining five predicted it as 'Recurrent'. The first column of table 5 shows the 'Left hand side rules (antecedent)'. Here are the input

Table 3 The detailed ROC analysis results for significant quantitative features

Feature	Cut-off value	Metrics with 95% CI			P value*
		AUC	Sensitivity	Specificity	
Serum sodium	≤134.5 (mEq/L)	0.972 (0.948 to 0.995)	0.875 (0.71 to 0.965)	0.932 (0.847 to 0.977)	**<0.001**
Body temperature	≤38.65 (°C)	0.793 (0.699 to 0.886)	0.656 (0.468 to 0.814)	0.781 (0.669 to 0.869)	**<0.001**
Haemoglobin	≤123.5 (g/L)	0.634 (0.518 to 0.749)	0.938 (0.792 to 0.992)	0.343 (0.235 to 0.463)	**0.0234**
Calculated osmotic pressure	≤282.57 (mOsm/kg water)	0.838 (0.746 to 0.930)	0.75 (0.566 to 0.885)	0.863 (0.762 to 0.932)	**<0.001**
Time from onset of fever to seizure	≤7.5 (hours)	0.965 (0.926 to 1.000)	0.875 (0.71 to 0.965)	0.986 (0.926 to 1.000)	**<0.001**
Time from onset of seizure to arrival at the ED	≤35.5 (min)	0.711 (0.603 to 0.820)	0.656 (0.468 to 0.814)	0.712 (0.594 to 0.812)	**<0.001**

P values which are≤0.05 are reported in bold.
*Calculated p value for AUC.
AUC, area under the curve; ED, emergency department; ROC, receiver operating characteristics.

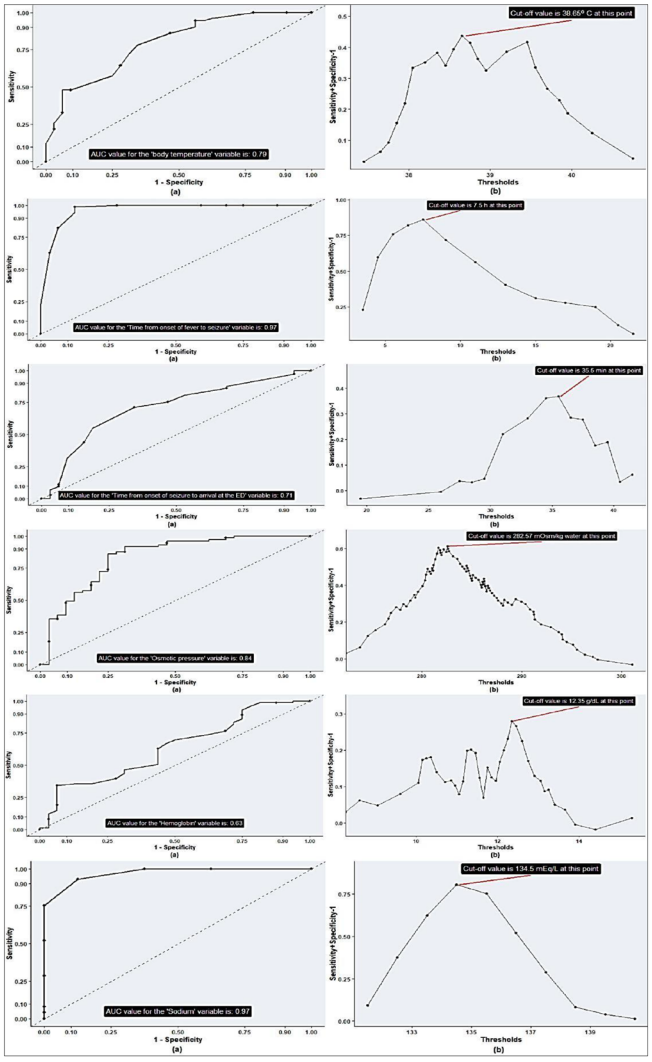

Figure 1 Area under the ROC curve analysis results (a), Youden index variations by cut-off values (b). AUC, area under the curve; ED, emergency department; ROC, receiver operating characteristics.

rules formed by the combination of features discretised according to an optimal cut-off point with the CACC discretisation method. The second column of table 5 shows the 'Right hand side rules (consequent)'. This column shows the classification output (SFS or RFS) produced from the rules in the first column. For easy readability of table 5, feature names are abbreviated and explanations of abbreviations are inserted as footnotes to table 5. According to table 5, the classification rule with the highest support value was:

IF time from onset of fever to seizure = (3–7.5 hours) and sodium = (134–141 mEq/L) than FS type = SFS

The confidence value for this rule was 0.985. This means that when the rule 'time from onset of fever to seizure' = (3, 7.5) AND sodium = (134, 141) is satisfied, the conditional probability that the FS type is classified as 'Simple' is 0.985. The antecedent rule, had the highest support value among the rules whose consequent rule was obtained as 'RFS'. The confidence value for the relevant classification rule is calculated as 1.

IF time from onset of fever to seizure = (7.5–22 hours) and sodium = (131–134 mEq/L) than FS type = RFS

In online supplemental figure 2, the network visualisation of classification-based association rules generated by CBA is presented.

Based on the learnt classification rules, FS type was predicted for each instance by applying the CBA model to the entire dataset. The results are shown in table 6. According to table 6, the CBA model correctly classified 102 instances and misclassified 3 instances. The metrics used to measure the classification performance of the CBA model and the values calculated for these metrics are given in table 7.

When tables 6 and 7 are considered together, it is observed that the CBA model misclassified 2 out of 105 samples, resulting in a model classification accuracy of 97%. Three observations were predicted as SFS by the model although they were actually classified as RFS. Therefore, the sensitivity value of the model was calculated as 91%. Since all SFS samples were correctly classified, the specificity value was calculated as 100%.

DISCUSSION

The study aimed to develop a CBA model that predicts the risk of developing RFS in children with FSs. Univariate analyses identified potential risk factors associated with the development of RFS, including seizure duration, number of recurrent seizures within 24 hours, family history, body temperature, time from fever onset to seizure, time from seizure onset to ED arrival, hyponatraemia, osmotic pressure and low haemoglobin. A total of 11 classification rules were obtained for the two patient groups as a result of the application of the CBA algorithm. The CBA model achieved 97% accuracy in

Table 4 Multiple logistic regression analysis results

Goodness of fit statistics			Model coefficient's statistics			
Hosmer-Lemeshow test			Feature	β	P value	OR (95% CI)
χ^2	*df*	*p*	Body temperature (°C)	–2.866	**0.004**	0.057 (0.008 to 0.397)
0.908	7	0.996	Serum sodium (mEq/L)	–3.156	**0.002**	0.043 (0.006 to 0.301)

P values which are ≤0.05 are reported in bold.
df, degree of freedom; β, coefficient.

Table 5 The generated classification rules by CBA algorithm

Left hand side rules (antecedent)	Right hand side rule (consequent)	Support	Confidence	Frequency of the rule
TFS=(3, 7.5) and Na=(134, 141)	Simple febrile seizure	0.629	0.985	66
TFS=(3, 7.5) and TSED=(14, 35.5) and Na=(134, 141)	Simple febrile seizure	0.457	1	48
TFS=(3, 7.5) and Hb=(101, 119) and Na=(134, 141)	Simple febrile seizure	0.381	1	40
BT=(39.5, 41) and TFS=(3, 7.5)	Simple febrile seizure	0.295	1	31
BT=(39.5, 41) and Na=(134, 141)	Simple febrile seizure	0.276	1	29
TFS=(3, 7.5) and Hb=(124, 158)	Simple febrile seizure	0.219	1	23
TFS=(7.5, 22) and Na=(131, 134)	Recurrent febrile seizure	0.219	1	23
TFS=(7.5, 22) and TSED=(35.5, 42)	Recurrent febrile seizure	0.162	1	17
TFS=(7.5, 22) and Hb=(101, 119)	Recurrent febrile seizure	0.124	1	13
BT=(37.3, 38.3) and TFS=(7.5, 22)	Recurrent febrile seizure	0.114	1	12
BT=(37.3, 38.3) and Na=(131, 134)	Recurrent febrile seizure	0.105	1	11

BT, body temperature (°C); CBA, Classification Based on Association Rules Algorithm; Hb, haemoglobin (g/L); Na, sodium (mEq/L); TFS, time from onset of fever to seizure (hours); TSED, time from onset of seizure to arrival at the emergency department (min).

overall classification performance, which shows that CBA has good predictive ability. Well-defined rules derived from this model may provide clinicians with an easily interpretable/applicable decision support system to estimate the risk of RFS.

In the present study, multiple predictors were identified indicating the multifactorial aetiology of RFS. Considering the results of univariate analyses, a statistically significant difference was observed among FS subtypes for seizure duration, number of seizures, family history, body temperature, time from fever onset to seizure, time from seizure onset to arrival at the ED, calculated osmotic pressure, haemoglobin and sodium variables. ROC analysis was applied to measure the ability of quantitative variables to distinguish FS groups and to determine threshold values for the variables in question. The results of the ROC analysis showed that the calculated AUC values were statistically significant. Serum sodium and time from fever onset to seizure variables have the two highest AUC values. Looking at the MLRA results, the number of predictors included in the model is only two. This is because of the stepwise model selection approach used to improve the predictive performance of the model. When the findings obtained here were evaluated, it was seen that each one-unit increase in body temperature and serum sodium predictors reduced the risk of RFS by 83.5% (OR=0.057, p=0.004) and 76.8% (OR=0.043, p=0.002), respectively.

The CBA model was used to classify FS subtypes and obtain interpretable rule patterns. The CBA model determined six rules for SFS and five rules for RFS to classify FS subtypes based on the data. Considering the classification performance obtained from these rules, it was seen that the CBA model misclassified only three children with SFS. When the rules produced to predict RFS were evaluated, it was seen that the rule patterns mostly included the time from fever onset to seizure, the time from seizure onset to arrival at the ED, serum sodium and body temperature variables. The rules obtained for these estimators include certain ranges. In general, the time from fever onset to seizure is 7.5–22 hours, serum sodium level is 131–134 mEq/L and body temperature is 37.3–38.3°C. If the time from the beginning of the seizure to arrival at the ED is over 35.5 min, the confidence values obtained are seen to be 1. This means that when these rules come

Table 6 The confusion matrix for the CBA model

	Actual classes		
Model prediction	Simple febrile seizure	Recurrent febrile seizure	Total
Simple febrile Seizure	73	3	76
Recurrent febrile Seizure	0	29	29
Total	73	32	105

CBA, Classification Based on Association Rules Algorithm.

Table 7 The classification performance of the CBA model

Metric	Value (95% CI)
Accuracy	0.97 (0.92 to 0.99)
Sensitivity	0.91 (0.75 to 0.98)
Specificity	1.00 (0.95 to 1.00)

CBA, Classification Based on Association Rules Algorithm.

together, RFS is estimated with 100% probability. When the frequencies of the rule patterns are taken into consideration, it is seen that the most repeated rule is the rule involving the time from the onset of fever to the seizure and the serum sodium level. Considering the haemoglobin rule, the range determined for the relevant variable (101–119 g/L) is in both SFS and RFS. When the relevant rule regarding the time from seizure onset to arrival at the ED was examined, the time from fever onset to arrival at the seizure was less than 35.5 min, which had a high confidence value in predicting SFS. When the rules on the RFS side were considered, it was seen that the time from seizure onset to arrival at the ED was calculated as over 35.5 min. This finding allows us to conclude that it is completely consistent with the threshold value obtained in the ROC analysis for the time from seizure onset to arrival at the ED.

It can be seen that in the classification rules obtained from the CBA model, cut-off points are also determined for the variables (table 5). Although these cut-off points are compatible with the cut-off points obtained by ROC analysis, small differences are observed. This is because the CBA model takes other predictor variables into account when creating the prediction rules. In this context, it can be argued that the cut-off points obtained from the CBA model are more valuable than the cut-off points obtained from the single ROC analysis.

In the current study, the incidence of RFS was found to be 30.5%. This rate is slightly higher than that reported in the literature.[5–12] Previous studies have identified several predictors of RFS, including family history of FS,[10 15] male gender,[11] history of recurrent FS,[12] time from fever onset to seizure,[12] lower body temperature,[8 11 12] lower C-reactive protein and blood glucose levels,[8] hyponatraemia,[35 36] non-generalised seizure types and prolonged seizure duration.[37] However, inconsistencies have been observed in these studies. This may be due to the selection of different sampling frames in determining the cohorts used in the studies, as well as the diversity in the univariate and multivariate analysis methods used.[12]

While hyponatraemia was confirmed as a risk factor for RFS in the present study,[35 36] several studies have reported that serum sodium levels do not predict recurrence of FS.[10 38] The development of hyponatraemia may be due to a slight increase in antidiuretic hormone during acute fever, infection and resulting fluid retention; relative hyponatraemia may lower the fluid threshold, exacerbate hyponatraemia and trigger FS.[35] This suggests that it is not possible to conclude whether RFS affects hyponatraemia or hyponatraemia affects RFS. Kubota *et al*[29] found that the calculated osmotic pressure of children with prolonged seizures was significantly higher than that of children with simple and multiple seizures during the same febrile illness. Andrew[39] found that lower osmotic pressure contributed to the onset of seizures. The results of this study suggest that prolonged seizures may be triggered by other mechanisms. Large multicentre studies are needed to evaluate the relationship between osmotic pressure and FS. Similar to our study, easily measured inflammatory markers, including NLR, were not useful for distinguishing between types of FS in children.[29] In the current study, the risk of RFS was significantly higher at lower body temperature, confirming the findings of a previous study suggesting that a lower temperature may indicate a lower seizure threshold in individuals with genetically predisposed brains.[40] In addition, when there is a history of early recurrence or when the seizure occurs with low body temperature, it is recommended that the patient be closely monitored for 6–8 hours.[12]

Limitations

This study has several limitations. First, it is retrospective, has a relatively small sample size, and is based on a dataset collected from a single clinical centre. Second, we did not investigate the long-term neurological outcomes of patients with RFS, including the subsequent development of epilepsy, cognitive impairment or psychiatric disorders. However, FS recurrence is not expected to have a significant impact on long-term prognosis.[18] Finally, external validation with datasets from different clinical centres is necessary to validate the findings obtained from the CBA model.

CONCLUSION

In this study, we developed a CBA model with good predictive performance for RFS. This model can be used as a risk estimator to identify children at high risk of seizure recurrence. Additionally, well-defined rules obtained from the model can provide a decision support system for clinicians. The established rules can be used to generate hypotheses for subsequent studies, which can improve evidence-based clinical studies.

Contributors Conception/design of the study: GY, AKA, BÖ. Data collection: GY, BÖ. Data analysis and interpretation: GY, AKA, BÖ, SG. Drafting the article: GY, AKA. Critical revision of the article: BÖ, SG. All authors: final approval of the version to be published. Guarantors: GY, AKA, BÖ, SG are responsible for the overall consent (as guarantors)

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

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

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by the ethic committee of the Medical Faculty, Inönü University (Ethics Approval number 2023/4505). Since this study was designed as a retrospective cohort study, informed consent was not obtained from the participants.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID iD

Gül Yücel http://orcid.org/0000-0001-5753-9048

REFERENCES

1 Steering Committee on Quality Improvement and Management, Subcommittee on Febrile Seizures American Academy of Pediatrics. Febrile seizures: clinical practice guideline for the long-term management of the child with simple febrile seizures. *Pediatrics* 2008;121:1281–6.
2 Tsuboi T. Epidemiology of febrile and afebrile convulsions in children in Japan. *Neurology (ECronicon)* 1984;34:175–81.
3 Sugai K. Current management of febrile seizures in Japan: an overview. *Brain Dev* 2010;32:64–70.
4 Byeon JH, Kim GH, Eun BL. Prevalence, Incidence, and Recurrence of Febrile Seizures in Korean Children Based on National Registry Data. *J Clin Neurol* 2018;14:43–7.
5 Henry C, Cockburn C, Simpson MH, *et al*. The baseline risk of multiple febrile seizures in the same febrile illness: a meta-analysis. *Eur J Pediatr* 2022;181:2201–13.
6 Patel AD, Vidaurre J. Complex febrile seizures: a practical guide to evaluation and treatment. *J Child Neurol* 2013;28:762–7.
7 Inoue M, Adachi S, Kawakami I, *et al*. Change in the strategy for prophylactic diazepam use for febrile seizures and the impact on seizure recurrence within 24 h. *Seizure* 2020;75:70–4.
8 Kubota J, Higurashi N, Hirano D, *et al*. Body temperature predicts recurrent febrile seizures in the same febrile illness. *Brain Dev* 2021;43:768–74.
9 Murata S, Okasora K, Tanabe T, *et al*. Acetaminophen and Febrile Seizure Recurrences During the Same Fever Episode. *Pediatrics* 2018;142:e20181009.
10 Maksikharin A, Prommalikit O. Serum sodium levels do not predict recurrence of febrile seizures within 24 hours. *Paediatr Int Child Health* 2015;35:44–6.
11 Kubota J, Higurashi N, Hirano D, *et al*. Predictors of recurrent febrile seizures during the same febrile illness in children with febrile seizures. *J Neurol Sci* 2020;411:116682.
12 Kim JS, Woo H, Kim WS, *et al*. Clinical Profile and Predictors of Recurrent Simple Febrile Seizure. *Pediatr Neurol* 2024;156:4–9.
13 Subcommittee on Febrile Seizures, American Academy of Pediatrics. Neurodiagnostic evaluation of the child with a simple febrile seizure. *Pediatrics* 2011;127:389–94.
14 Raghavan VR, Porter JJ, Neuman MI, *et al*. Trends in Management of Simple Febrile Seizures at US Children's Hospitals. *Pediatrics* 2021;148:e2021051517.
15 Guedj R, Chappuy H, Titomanlio L, *et al*. Do All Children Who Present With a Complex Febrile Seizure Need a Lumbar Puncture? *Ann Emerg Med* 2017;70:52–62.
16 Olson H, Rudloe T, Loddenkemper T, *et al*. Should patients with complex febrile seizure be admitted for further management? *Am J Emerg Med* 2018;36:1386–90.
17 Kajiwara K, Koga H. Risk factors for acute encephalitis and early seizure recurrence in complex febrile seizures. *Eur J Pediatr* 2022;181:3103–10.
18 Mewasingh LD, Chin RFM, Scott RC. Current understanding of febrile seizures and their long-term outcomes. *Dev Med Child Neurol* 2020;62:1245–9.
19 Verity CM, Golding J. Risk of epilepsy after febrile convulsions: a national cohort study. *BMJ* 1991;303:1373–6.
20 Sutton RT, Pincock D, Baumgart DC, *et al*. An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ Digit Med* 2020;3:17.
21 Soni J, Ansari U, Sharma D, *et al*. Predictive data mining for medical diagnosis: an overview of heart disease prediction. *IJCA* 2011;17:43–8.
22 Pazhanikumar K, Arumugaperumal S. Association rule mining and medical application: a detailed survey. *IJCA* 2013;80:10–9.
23 Piran N, Farhadian M, Soltanian AR, *et al*. Diabetic foot ulcers risk prediction in patients with type 2 diabetes using classifier based on associations rule mining. *Sci Rep* 2024;14:635.
24 Zhang W, Ma D, Yao W. Medical diagnosis data mining based on improved apriori algorithm. *JNW* 2014;9:1339.
25 Das PP, Mast M, Wiese L, *et al*. Data extraction for associative classification using mined rules in pediatric intensive care data. 2023.
26 Alwidian J, Elhassan A, Ghnemat R, *et al*. Predicting autism spectrum disorder using machine learning technique. *IJRTE* 2020;8:4139–43.
27 Jabez Christopher J, Khanna Nehemiah H, Kannan A. A clinical decision support system for diagnosis of Allergic Rhinitis based on intradermal skin tests. *Comput Biol Med* 2015;65:76–84.
28 Capovilla G, Mastrangelo M, Romeo A, *et al*. Recommendations for the management of "febrile seizures": Ad Hoc Task Force of LICE Guidelines Commission. *Epilepsia* 2009;50 Suppl 1:2–6.
29 Kubota J, Hirano D, Suzuki T, *et al*. The role of inflammatory markers and calculated osmotic pressure in the classification of febrile seizures. *Eur Rev Med Pharmacol Sci* 2020;24:11187–91.
30 Arslan AK, Yaşar Ş, Çolak C, *et al*. WSSPAS: an interactive web application for sample size and power analysis with R using Shiny. *Turkiye Klinikleri J Biostat* 2018;10:224–46.
31 Liu B, Hsu W, Ma Y. Integrating classification and association rule mining. Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining; 1998:80–6.
32 Agrawal R, Srikant R. Fast algorithms for mining association rules. Proc. 20th Int. Conf. Very Large Data Bases, VLDB; 1994:487–99.
33 Tsai CJ, Lee CI, Yang WP. A discretization algorithm based on class-attribute contingency coefficient. *Inf Sci (Ny)* 2008;178:714–31.
34 Hicks SA, Strümke I, Thambawita V, *et al*. On evaluation metrics for medical applications of artificial intelligence. *Sci Rep* 2022;12:5979.
35 Hugen CA, Oudesluys-Murphy AM, Hop WC. Serum sodium levels and probability of recurrent febrile convulsions. *Eur J Pediatr* 1995;154:403–5.
36 Kiviranta T, Airaksinen EM. Low sodium levels in serum are associated with subsequent febrile seizures. *Acta Paediatr* 1995;84:1372–4.
37 Jeong JH, Lee JH, Kim K, *et al*. Rate of and risk factors for early recurrence in patients with febrile seizures. *Pediatr Emerg Care* 2014;30:540–5.
38 Thoman JE, Duffner PK, Shucard JL. Do serum sodium levels predict febrile seizure recurrence within 24 hours? *Pediatr Neurol* 2004;31:342–4.
39 Andrew RD. Seizure and acute osmotic change: clinical and neurophysiological aspects. *J Neurol Sci* 1991;101:7–18.
40 Smith DK, Sadler KP, Benedum M. Febrile Seizures: Risks, Evaluation, and Prognosis. *Am Fam Physician* 2019;99:445–50.