Pediatric Surgery International (2023) 39:9
https://doi.org/10.1007/s00383-022-05309-6

ORIGINAL ARTICLE

Predicting recurrent cases of intussusception in children after air enema reduction with machine learning models

Jing-yan Guo[1] · **Yu-feng Qian**[2]

Accepted: 11 November 2022

Abstract

Purpose To develop a model to identify risk factors and predict recurrent cases of intussusception in children.

Methods Consecutive cases and recurrent cases of intussusception in children from January 2016 to April 2022 were screened. The cohort was divided randomly at a 4:1 ratio to a training dataset and a validation dataset. Three parallel models were developed using extreme gradient boosting (XGBoost), logistic regression (LR), and support vector machine (SVM). Model performance was assessed by the area under the receiver operating characteristic curves (AUC).

Results A total of 2469 cases of intussusception were included, where 225 were recurrent cases. The XGBoost (AUC = 0.718) models showed the best performance in the validation dataset, followed by the LR model (AUC = 0.652), while the SVM model (AUC = 0.613) performed worst among the three models. Based on the Shapley Additive exPlanation values, the most important variables in the XGBoost models were air enema pressure, mass size, age, duration of symptoms, and absence of vomiting.

Conclusions Machine learning models, especially XGBoost, could be used to predict recurrent cases of intussusception in children. The most important contributing factors to the models are air enema pressure, mass size, age, duration of symptoms and absence of vomiting.

Keywords Intussusception · Recurrence · Child · Risk factor · Machine learning

Introduction

Recurrent intussusception is not uncommon in clinical practice and often manifests as acute bowel obstruction. Previous studies have noted that the rate of recurrence of intussusception ranges from 8 to 20% [1–3]. Recurrent intussusception is a challenge for most pediatricians, radiologists, and pediatric surgeons because it is not predictable and preventable. When recurrent epitope happens, it can compromise the blood supply of the bowels [4]. Any delay in proper diagnosis and treatment of such cases may lead to serious complications, such as bowel necrosis and perforation, which may even require intestinal resection [1]. Timely diagnosis and treatment of recurrent intussusception are vital to avoid such serious complications. Understanding the factors associated with the risk of recurrent intussusception is crucial for early diagnosis and treatment of recurrent intussusception.

Previous studies have identified some of the factors associated with the increased risk of recurrent intussusception [5, 6]. However, the sample size was not sufficiently large; so, we still cannot identify a reliable risk factor, and it is unclear which factors are more important for the recurrence of intussusception. The reasons behind the differing risk of recurrent intussusception are still not clear. In recent years, machine learning models have been widely used in medical practice for diagnosis, treatment, and decision- making in various diseases [7, 8]. To date, there have been no related machine learning models to predict recurrent intussusception. The aim of the present study was to develop and validate a machine learning model to predict recurrent intussusception and identify risk factors for recurrent intussusception.

✉ Yu-feng Qian
qianyufeng95@163.com

1 School of Health Economics and Management, Nanjing University of Chinese Medicine, Nanjing 210023, China

2 Department of Radiology, Children's Hospital of Soochow University, Suzhou 215025, China

Published online: 28 November 2022

Methods

Study population

An over 5 year retrospective cohort study was conducted on pediatric patients who underwent air enema or surgery for intussusception between January 2016 and April 2022 at the Children's Hospital of Soochow University. The inclusion criteria were as follows: intussusceptions were first diagnosed by ultrasound (US) and then treated with air enema reductions. All of the patients first underwent air enema reduction. In case of failure of air enema reduction, treatment by surgical reduction was applied. Patients with symptoms and diagnosis by US, which were later not confirmed by air enema reduction, were excluded. Recurrent intussusceptions were defined as recurrence after a successful air enema reduction or surgical reduction. Nonrecurrent intussusception was defined as successful reduction after air enema or surgery without recurrence. Multiple-recurrent intussusceptions were defined as two or more recurrences after successful air enema reductions or surgery [5, 9].

Data collection

Clinical, imaging, and pathological data were collected, including sex, age, weight, date of episode, abdominal pain, vomiting, paroxysmal crying, bloody stool, mass, duration of symptoms, diarrhea, episode times, mass size on US, intussusception intestinal effusion on US, air enema pressure, mass location, pathological lead points (PLPs), and enlarged lymph nodes. Mass location was defined as the right abdomen or left abdomen depending on whether the mass was located on the right or on the left of the spine, respectively. Enlarged lymph nodes were defined as those with a diameter greater than 15 mm. Mass size was defined as the long diameter × short diameter in the maximum section of US. Each patient was followed-up with for at least 6 months. Informed consent was obtained from the designated guardians of each patient. The present study was approved by the Ethics Committees of Children's Hospital of Soochow University and conformed to the Declaration of Helsinki.

Statistical analysis

Statistical analysis was performed with Python (version 3.9), with p values less than 0.05 considered statistically significant. Data were presented as numbers (n) and percentages. Univariate comparisons were made using nonparametric one-way Wilcoxon's rank-sum, $\chi 2$, and Fisher's exact tests depending on the statistical distribution. For continuous variables, we used the Student's t test or Kruskal–Wallis test, depending on the statistical distribution. The data were randomly divided, with 80% used for training and 20% for validation. Three parallel classifiers were developed based on eXtreme Gradient Boosting (XGBoost), support vector machine (SVM), and logistic regression (LR) to develop appropriate models and identify the factors that contribute to recurrent intussusception after air enema reduction or surgery. The performance of various developed classifiers was assessed by receiver operating characteristic (ROC) analysis. Shapley Additive explanation (SHAP) values were used to provide consistent and locally accurate attribution values for each feature within each prediction model and to evaluate the importance of individual features [8, 10]. Finally, clinical impact curve (CIC) nomogram was plotted to evaluate the clinical usefulness and applicability net benefits of the model.

Results

Clinical characteristics of recurrent cases of intussusception

A total of 2469 pediatric patients with intussusception were admitted to our hospital (Fig. 1). There were 1570 male patients and 899 female patients. A total of 1803 patients were less than 1 year old. A total of 225 cases were recurrent cases of intussusception with 626 episodes of recurrence intussusception. Among the recurrent cases, 140 had one episode of recurrence, 42 had two recurrences, 21 had three recurrences, and 22 had more than four recurrences (Fig. 2). Recurrence occurred within 24 h in 163 cases, between 24 h and 1 week in 85 cases, between 1 week and 1 month in 11 cases, between 1 and 6 months in 45 cases, between 6 months and 1 year in 32 cases, and after more than 1 year in 44 cases (Fig. 3). The age of the patients with recurrence was from 4 months to 12 years, with a mean age of 2.0 years (median 2.0 [730 days]; range 0–12 years). There were 152 male patients and 73 female patients. Fifty-one patients were in the younger group (< 1 year), and 174 patients were in the older group (≥ 1 year). The main clinical symptoms included paroxysmal crying ($n = 32$), abdominal pain ($n = 116$), and vomiting ($n = 44$). Five patients had bloody stool and five had diarrhea. Eighteen patients had PLPs (duplication: $n = 8$; polyp: $n = 4$; Meckel's diverticulum, $n = 2$; tumor: $n = 4$). Four patients had intestinal necrosis and underwent intestinal resection, of which one patient experienced perforation. Ten patients with intussusception had intestinal effusion, and 29 patients had enlarged lymph nodes on US. The duration of symptoms was less than 12 h in 171 cases in recurrent intussusception.

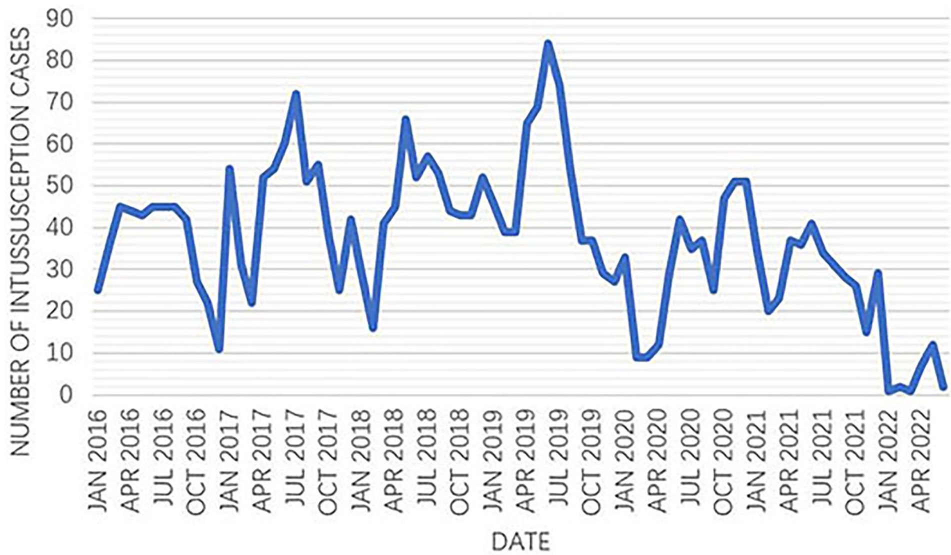

Fig. 1 Monthly observed morbidity of intussusception from January 2016 to April 2022 in Suzhou

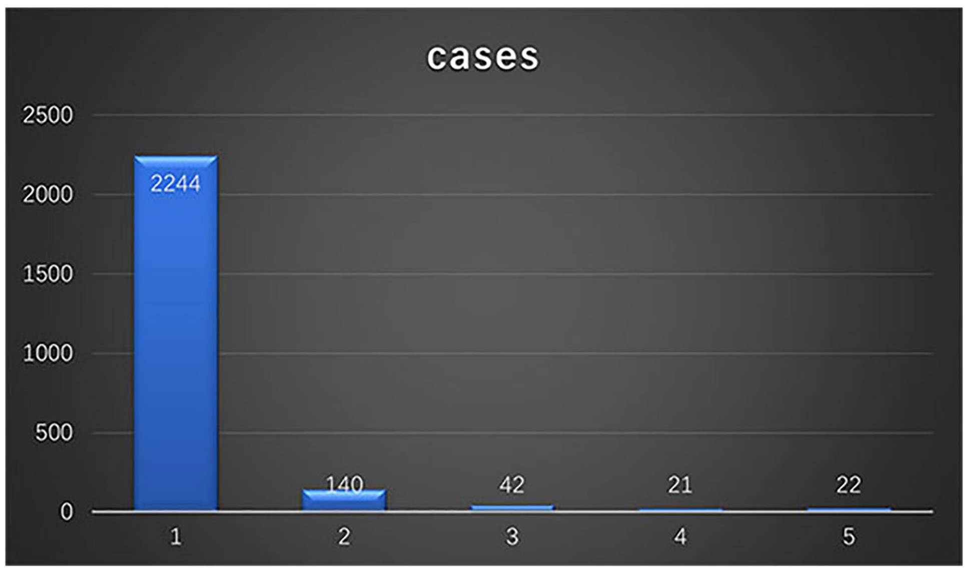

Fig. 2 Distribution of episode times of intussusceptions

Predictors of recurrent intussusception

Table 1 shows the demographic and clinical features of the training and validation datasets. Nineteen features that are readily available in routine practice, namely, air enema pressure, intussusception size on US, age, duration of symptoms, vomiting, paroxysmal crying, abdominal pain, mass, mass location, sex, enlarged lymph nodes, fever, PLPs, bloody stool, intussusception intestinal effusion on US, intussusception shape, abdominal distention, diarrhea, and constipation, were used in the model development.

Machine learning model analysis

The XGBoost showed the best performance among the three models, with an AUC of 0.778 (95% CI 0.737, 0.819) in the training set and 0.718 (95% CI 0.631, 0.806) in the test set. The LR model showed an AUC of 0.657(95% CI 0.612, 0.702) in the training set and 0.638 (95% CI 0.547, 0.728) in the test set (Fig. 4 and Table 2). SVM had the poorest performance among the three models, with an AUC of 0.619 (95% CI 0.573, 0.664) in the training set and 0.555 (95% CI 0.465, 0.646) in the test set. The sensitivity and specificity of the XGBoost model at optimal cut-off in the validation were

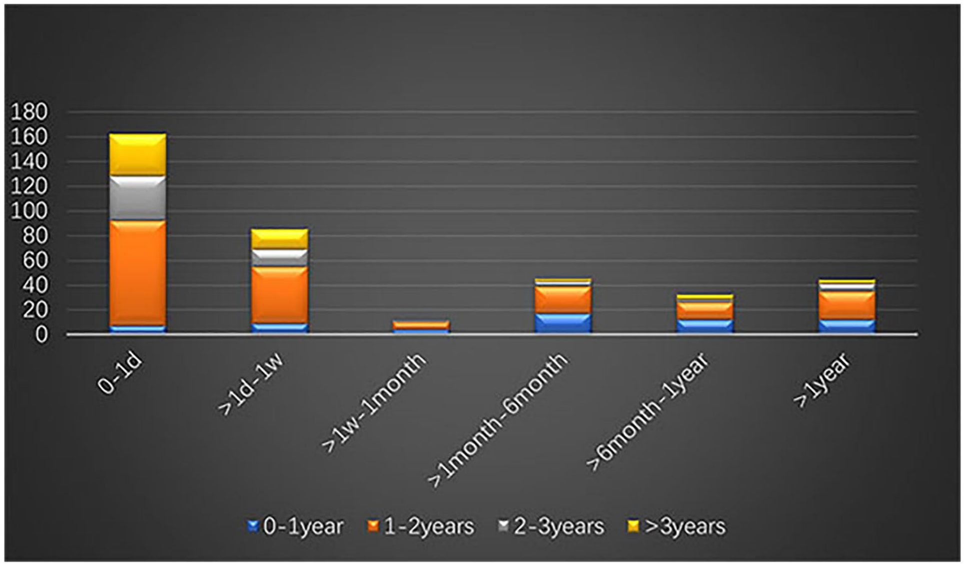

Fig. 3 Distribution of recurrent episode time points among four age groups

Table 1 Demographic and clinical features of the training and validation datasets

Variables	All (2469)	Training set (1975)	Test set (494)	*p* value
Sex (male %)	1570 (63.6)	1261 (63.8)	309 (62.6)	0.592
Age (day)	630.0 (360.0–730.0)	630.0 (360.0–730.0)	570.0 (360.0–730.0)	0.000
Abdominal pain (%)	1181 (47.8)	950 (48.1)	231 (46.8)	0.594
Paroxysmal crying (%)	295 (11.9)	239 (12.1)	56 (11.3)	0.639
Vomiting (%)	652 (26.4)	501 (25.4)	151 (30.6)	0.019
Bloody stool (%)	65 (2.6)	50 (2.5)	15 (3.0)	0.531
Mass (%)	7 (0.3)	6 (0.3)	1 (0.2)	0.705
Constipation (%)	15 (0.6)	11 (0.6)	4 (0.8)	0.518
Fever (%)	58 (2.3)	45 (2.3)	13 (2.6)	0.643
Abdominal distention (%)	1 (0.0)	1 (0.1)	0 (0.0)	0.617
Diarrhea (%)	41 (1.7)	34 (1.7)	7 (1.4)	0.636
Duration of symptoms	1231.407 (720.0–1231.407)	1231.407 (720.0–1231.4)	1231.407 (720.0–1231.4)	0.000
intussusception size in US	1088.0 (930.0–1248.0)	1088.0 (929.0–1248.0)	1082.5 (930.0–1224.0)	0.000
Enlarged lymph nodes (%)	188 (7.6)	152 (7.7)	36 (7.3)	0.759
Intussuscepiton shape (%)	1 (0.0)	1 (0.1)	0 (0.0)	0.617
Intussuscepiton intestinal effusion in US (%)	60 (2.4)	46 (2.3)	14 (2.8)	0.515
Air enema pressure	8.0 (8.0–10.0)	8.0 (8.0–10.0)	8.0 (8.0–10.0)	0.000
Mass location (right abdomen1) (%)	2397 (97.1)	1915 (97.0)	482 (97.6)	0.472
Pathological lead points (%)	8 (0.3)	5 (0.3)	3 (0.6)	0.216
Intestinal duplication (%)	8 (0.3)	8 (0.4)	0 (0.0)	0.157
Polyp (%)	4 (0.2)	2 (0.1)	2 (0.4)	0.134
Meckel's diverticulum (%)	2 (0.1)	2 (0.1)	0 (0.0)	0.479
Tumor (%)	4 (0.2)	2 (0.1)	2 (0.4)	0.134
Intestinal necrosis (%)	2 (0.1)	2 (0.1)	0 (0.0)	0.479
Intestinal resection (%)	2 (0.1)	1 (0.1)	1 (0.2)	0.289
Intestinal perforation (%)	0 (0.0)	0 (0.0)	0 (0.0)	0.000

0.555 and 0.793, respectively. The sensitivity and specificity of the LR model at optimal cut-off in the validation were 0.556 and 0.710, respectively. The sensitivity and specificity of the SVM model were 0.355 and 0.731, respectively. Figure 5 shows the ranks of feature importance in each model. The features in the XGBoost model, as ranked by the SHAP

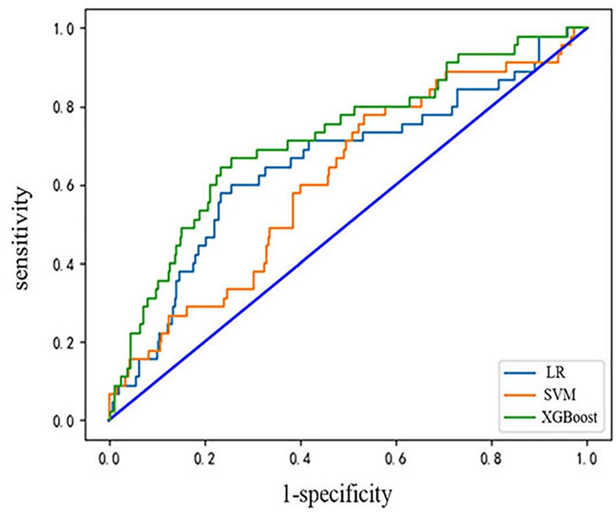

Fig. 4 Receiver operating characteristic (ROC) curves of the machine learning classifier. **A** XGBoost classifier, AUC = 0.778 and 0.718 in the training and validation datasets, respectively; **B** LR classifier, AUC = 0.657 and 0.638; **C** SVM classifier, AUC = 0.619 and 0.555

values, were air enema pressure, mass size on US, age, duration of symptoms, and vomiting (Fig. 6). To visualization of the XGBoost model, nomogram was made of 12 factors screened by XGBoost by logical regression algorithm (Fig. 7A). CIC analysis was performed in Fig. 7B to evaluate clinical applicability of the prediction nomogram. CIC analysis showed the potential clinical efficiency of the prediction model. When the threshold probability is over 20%, the high-risk population determined by the prediction model is highly matched with the actual population with recurrent cases intussusception.

Discussion

In this retrospective cohort study, we developed three models to predict the recurrence of intussusception in children using machine learning. The model based on XGBoost (AUC = 0.718) exhibited the best performance for the prediction of recurrent intussusception in children, compared with the LR model (AUC = 0.638) and the SVM model (AUC = 0.555). To the best of our knowledge, these models represent the first attempt to build mathematical models to predict recurrence of intussusception in children based on a large sample.

XGBoost is a state-of-art machine learning algorithm in prediction modeling. It has the capacity to analyze complex nonlinear relationships among various clinical factors [8, 11, 12]. XGBoost has the advantage of efficient and flexible processing of missing data, bootstrapping, and assembly of weak prediction models, which allows application in relatively small samples [13]. The LR model can accurately predict the probability of the binary dependent variable using maximum likelihood estimation to determine the regression coefficient [10]. SVM is a supervised learning model from statistical learning theory, and it has been widely used for

Table 2 Performance of the three machine learning models for predicting recurrent cases of intussusception in children

	Accuracy	Sensitivity	Specificity	AUC train (95% CI)	AUC test (95% CI)
Logistic regression	0.696 (344/494)	0.556 (25/45)	0.710 (319/449)	0.657 (0.612, 0.702)	0.638 (0.547, 0.728)
SVM	0.696 (344/494)	0.355 (16/45)	0.731 (328/449)	0.619 (0.573, 0.664)	0.555 (0.465, 0.646)
XGBoost	0.771 (381/494)	0.555 (25/45)	0.793 (356/449)	0.778 (0.737, 0.819)	0.718 (0.631, 0.806)

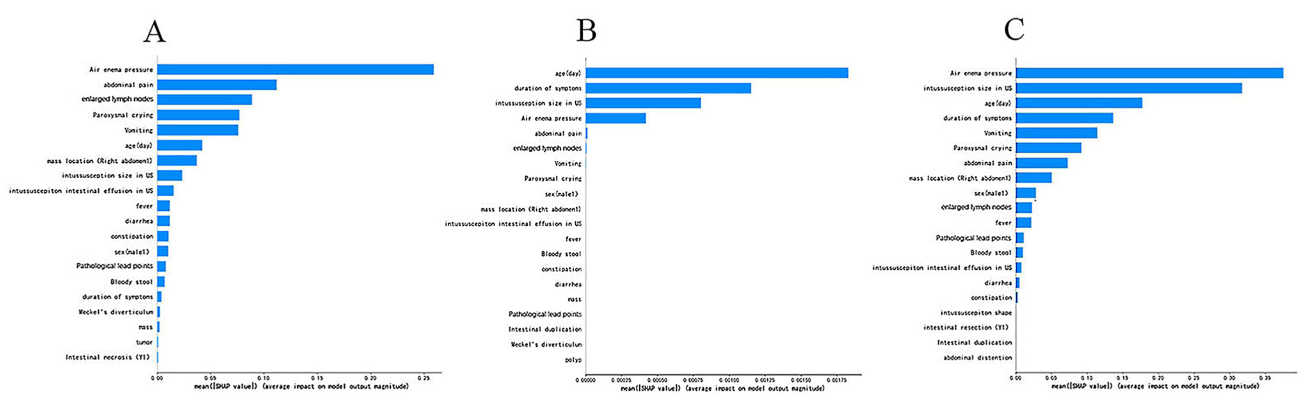

Fig. 5 Importance matrix plot of the machine learning models. **A** LR model; **B** SVM model; **C** XGBoost model

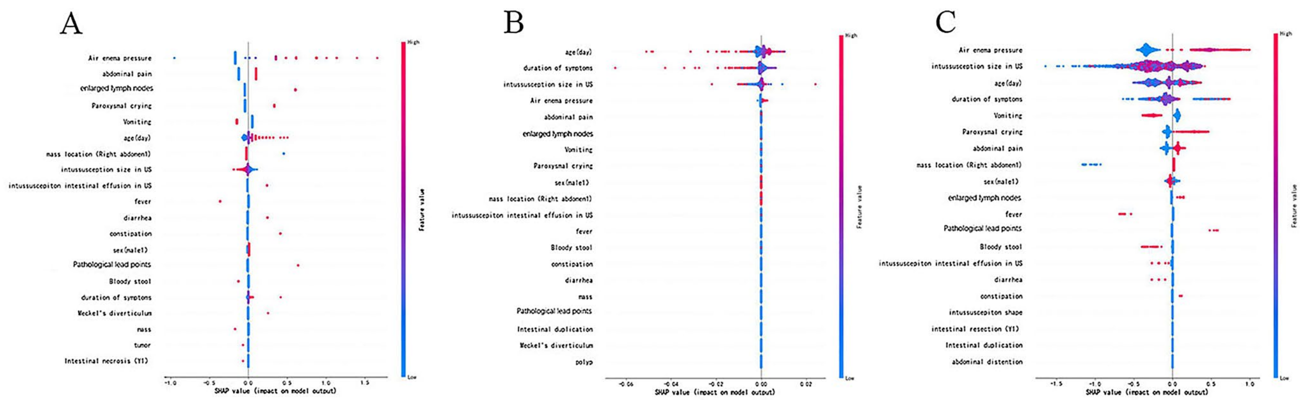

Fig. 6 SHAP (SHapley additive exPlanations) framework for the features in the three machine learning model (**A** = XGBoost model, **B** = LR model, **C** = SVM model). Input variables are ranked in descending order of feature importance. Red indicates high feature value; blue indicates low feature value

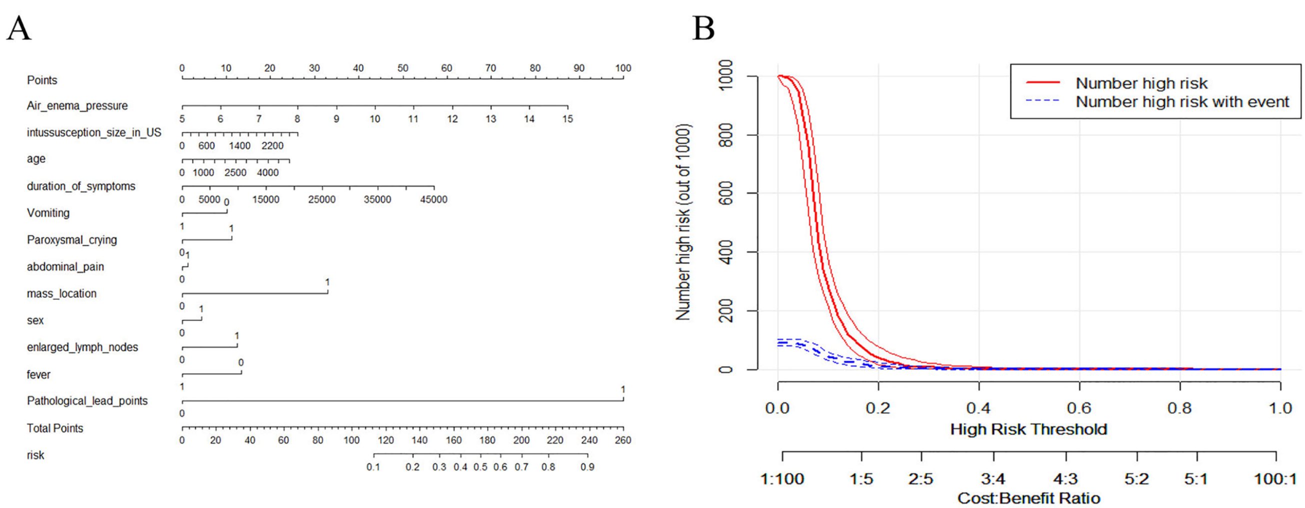

Fig. 7 **A** Nomogram to estimate the risk of recurrent cases of intussusception. firstly, a line with each parameter value to the score axis for the score, the points for all the parameters were then added, finally, a line from the total score axis was drawn to determine the risk of recurrent cases of intussusception. **B** Clinical impact curve (CIC) of XGBoost model. The red curve (number of high-risk individuals of recurrent) showed the number of cases who were classified as positive (high risk) by the model at each threshold probability; the blue curve (number of high-risk individuals with recurrent outcome) was the number of true positives at each threshold probability. CIC visually showed that nomogram conferred high clinical net benefit and confirmed the potential clinical value of the XGBoost model

its outstanding performance in pattern-recognition problems, even for small sample sizes. In addition, linear kernel-based SVM allows more features without easily overfitting [8, 14]. Consistent with previous studies on recurrent intussusception in general intussusception population [1], the XGBoost models outperformed the LR model and SVM in our study.

In the present study, a total of 20 variables were evaluated, but only 16 that are readily available in daily practice were included in model development to promote their potential use in clinical settings. These variables were air enema pressure, intussusception size on US, age, duration of symptoms, vomiting, paroxysmal crying, abdominal pain, mass location, sex, enlarged lymph nodes, fever, PLPs, bloody stool, intussusception intestinal effusion on US, diarrhea, and constipation. In the ROC analysis, the XGBoost model outperformed the LR model and the SVM model, with an AUC of 0.718 in the validation dataset. In addition to air enema pressure, intussusception size on US, age, duration of symptoms, and vomiting were the top features on the importance matrix plot, and the SHAP summary plot of all three models included variables that reflect intussusception itself.

The present study showed that air enema pressure in the first episode was the most important variable in predicting recurrent cases of intussusception in children. As shown in

Fig. 4, the risk of recurrent intussusception was higher when the air enema pressure was in the range of 8.7–12.0 kPa in the first episode, while an air enema pressure less than 8.7 kPa or greater than 12 kPa was associated with reduced risk of recurrence. The risk of recurrent intussusception increased with mass size in the range of 1053–1920 mm^2, while a mass size less than 1053 mm^2 or more than 1920 mm^2 was associated with reduced risk of recurrence. Zhang et al. reported that intussusception recurrence was associated with greater mass diameter (> 2.55 cm) [15]. Such a relationship between the mass size and air enema pressure has not been reported before and needs verification by future studies.

The current study showed that age was an important factor in predicting recurrence of intussusception. The risk of recurrent intussusception increased with the age in the range of 1080–1825 days, while age < 1080 or > 1825 days was associated with reduced risk of recurrence. Guo et al. reported 191 cases of recurrent intussusception and found that about 67.5% of the patients with recurrent intussusception were older than 1 year [5]. Chen et al. reported 1158 cases of recurrent intussusception and found that 77.5% of the patients with recurrent intussusception were older than 1 year [3]. These results are compatible with the current results that the old age group was more prone to recurrence after air enema reduction. The possible reason may involve some underlying factors such as the difference in thickness between the ileum and colon, prevalence of lead points, the number and volume of ileal lymphoid follicles, and ascending colon mobility, which all increase with age [2, 6, 16–18]. However, in the present study, we found that when the age was > 1825 days, the risk of recurrence decreased. Thus, the relationship with age should be further verified in future studies.

In the present study, we found that some clinical characteristics were associated with recurrence. Aabsence of vomiting, crying, abdominal pain, duration of symptoms and mass location (right abdomen) were also associated with recurrent intussusception. Guo and Justice et al. reported that the incidence of recurrent cases of intussusception was higher in patients with a mass location on the right side of the colon compared with those on the left side of the colon [5, 19]. Champoux et al. and Guo et al. reported that a shorter duration of intussusception was associated with recurrent intussusception. They also reported that the recurrent cases of intussusception often with absence of vomiting [5, 20]. The current study also found that patients with recurrent intussusception had a lower incidence of vomiting, which is similar to the above results. However, from the SHAP value of XGBoost, we cannot find the accurate pattern of relationship between recurrent cases and duration of symptoms, and which is different from the above studied. Besides that, we also found that patients with recurrent intussusception had a high rate of crying and abdominal pain, and the underlying reason should be further verified.

In addition to the above-mentioned clinical risk factors, we also found that patients with recurrent intussusception more often had enlarged lymph nodes on US. Chen et al. reported that in one of 22 patients, no other PLPs were identified, except for lymphoid hyperplasia [3]. Zhang et al. reported that intussusception recurrence was associated with enlarged abdominal lymph nodes [15], which is consistent with the present study. In the present study, PLPs were also a risk factor for recurrent intussusception. Xie et al. [6] and Lin et al. [21] reported that PLPs were a risk factor for intussusception recurrence, which is consistent with the present study. In such cases, once the PLP is detected on US, surgical intervention should be considered.

Limitations

The greatest limitations of the current study are its retrospective design and potential selection bias. Also, the models in the current study were not validated on external samples. Prospective and multicenter studies are needed to refine and optimize the prediction model in the future.

Conclusions

In conclusion, our algorithms based on XGBoost machine learning could be used to predict recurrent cases of intussusception in children. The most important factors contributing to the models are air enema pressure, intussusception size on US, age, duration of symptoms, and absence of vomiting. A prospective and multicenter study is warranted to verify and optimize our prediction model.

Author contributions Contributors Y-FQ designed the study; J-YG and Y-FQ collected data; J-YG analysed the data; J-YG and Y-FQ wrote and revised the manuscript; and all authors read and approved the final version of manuscript.

Declarations

Conflict of interest We have no conflicts of interest to declare.

Ethical approval The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study protocol was approved by the Ethics Committee of the Children's Hospital of Soochow University.

References

1. Niramis R, Watanatittan S, Kruatrachue A et al (2010) Management of recurrent intussusception: nonoperative or operative reduction? J Pediatr Surg 45(11):2175–2180
2. Cho MJ, Nam CW, Choi SH et al (2020) Management of recurrent ileocolic intussusception. J Pediatr Surg 55(10):2150–2153
3. Chen X, Chen Q, Wang X et al (2021) Clinical characteristics of recurrent intussusception: a single-center retrospective study. J Pediatr Surg 56(10):1831–1834
4. Guo WL, Wang J, Zhou M et al (2011) The role of plain radiography in assessing intussusception with vascular compromise in children. Arch Med Sci 7(5):877–881
5. Guo WL, Hu ZC, Tan YL et al (2017) Risk factors for recurrent intussusception in children: a retrospective cohort study. BMJ Open 7(11):e018604
6. Xie X, Wu Y, Wang Q et al (2018) Risk factors for recurrence of intussusception in pediatric patients: a retrospective study. J Pediatr Surg 53(11):2307–2311
7. Kourou K, Exarchos TP, Exarchos KP et al (2014) Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J13:8–17
8. Han X, Geng J, Zhang XX et al (2022) Using machine learning models to predict acute pancreatitis in children with pancreaticobiliary maljunction. Surg Today. https://doi.org/10.1007/s00595-022-02571-y
9. Ting X, Xufei D, Jiangbin L et al (2022) Development and validation of a nomogram for predicting pathological intussusceptions in children prior to surgical intervention. Front Pediatr 10:877358
10. Tseng PY, Chen YT, Wang CH et al (2020) Prediction of the development of acute kidney injury following cardiac surgery by machine learning. Crit Care 24(1):478
11. Goto T, Camargo CA Jr, Faridi MK et al (2018) Machine learning approaches for predicting disposition of asthma and COPD exacerbations in the ED. Am J Emerg Med 36:1650–1654
12. Goto T, Camargo CA Jr, Faridi MK et al (2019) Machine learning-based prediction of clinical outcomes for children during emergency department triage. JAMA Netw Open 2(1):e186937
13. Yuan KC, Tsai LW, Lee KH et al (2020) The development an artificial intelligence algorithm for early sepsis diagnosis in the intensive care unit. Int J Med Inform 141:104176
14. Xu L, Yang P, Liang W et al (2019) A radiomics approach based on support vector machine using MR images for preoperative lymph node status evaluation in intrahepatic cholangiocarcinoma. Theranostics 9(18):5374–5385
15. Zhang Y, Shao CC, Wei XL et al (2022) Ultrasound findings to predict risk of recurrence in pediatric intussusception after air enema reduction. J Ultrasound Med 41(5):1227–1235
16. Kim JH, Lee JS, Ryu JM et al (2018) Risk factors for recurrent intussusception after fluoroscopy–guided air enema. Pediatr Emerg Care 34(7):484–487
17. Shen G, Zhang C, Li J, Zhang J et al (2018) Risk factors for short-term recurrent intussusception and reduction failure after ultrasound-guided saline enema. Pediatr Surg Int 34(11):1225–1231
18. Wang Z, He QM, Zhang H et al (2015) Intussusception patients older than 1 year tend to have early re-currence after pneumatic enema reduction. Pediatr Surg Int 31(9):855–858
19. Justice FA, Nguyen LT, Tran SN et al (2011) Recurrent intussusception in infants. J Paediatr Child Health 47:802–805
20. Champoux AN, Del Beccaro MA, Nazar-Stewart V (1994) Recurrent intussusception. Risks and features. Arch Pediatr Adolesc Med 148:474–478
21. Lin XK, Xia QZ, Huang XZ et al (2017) Clinical characteristics of intussusception secondary to pathologic lead points in children: a single-center experience with 65 cases. Pediatr Surg Int 33(7):793–797