

# An early prediction model to identify neurological complications of childhood influenza: a random forest model

Suyun Li<sup>1</sup>, Weiqiang Xiao<sup>1</sup>, Huixian Li<sup>1</sup>, Dandan Hu<sup>1</sup>, Kuanrong Li<sup>1</sup>, Qinglian Chen<sup>1</sup>, Guangming Liu<sup>1</sup>, Haomei Yang<sup>2</sup>, Yongling Song<sup>1</sup>, Qiuyan Peng<sup>1</sup>, Qiang Wang<sup>2</sup>, Shuyao Ning<sup>1</sup>, Yumei Xiong<sup>1</sup>, Wencheng Ma<sup>1</sup>, Jun Shen<sup>1</sup>, Kelu Zheng<sup>1</sup>, Yan Hong<sup>2</sup>, Peiqing Li<sup>1</sup>, and Sida Yang<sup>1</sup>

<sup>1</sup>Affiliation not available

<sup>2</sup>Guangzhou Women and Children's Medical Center

August 22, 2020

## Abstract

**Background:** NeurologicalEarly prediction model for seizures in influenza complications of influenza are associated with high morbidity and mortality in children. The prognosis could be improved if early treatments are undertaken. **Objective:** To establish and validate an early prediction model to discriminate among neurological complications such as seizures, acute influenza virus-associated encephalitis (IAE), and acute necrotizing encephalopathy (ANE) in children with influenza. **Methods:** This was a retrospective single-center case-control study conducted at the Guangzhou Women and Children's Medical Center in Guangzhou (GWCMC), China, from November 2012 to January 2020. The random forest model was used to screen the characteristics and construct an early prediction model for convulsions, IAE, and ANE. **Results:** Of the 433 patients (294 male, 139 female; median age 2.8 (1.7,4.8) years), 278 (64.2%) had seizures, 106 (24.5%) had IAE, and 49 (11.3%) had ANE; 348 patients were in the training set and 85 in the validation set. When 10 variables were included, the cross-validation error was minimal; convulsions, procalcitonin, urea,  $\gamma$ -glutamyltransferase, aspartate aminotransferase, albumin/globulin ratio,  $\alpha$ -hydroxybutyric dehydrogenase, alanine aminotransferase, alkaline phosphatase, and C-reactive protein were included. The likelihood of having only seizures decreased with increasing procalcitonin, urea,  $\gamma$ -glutamyltransferase,  $\alpha$ -hydroxybutyric dehydrogenase, alanine aminotransferase, and aspartate aminotransferase, and with decreasing albumin/globulin ratio and alkaline phosphatase. The prediction model gave a prediction accuracy of 84.2%. **Conclusion:** This model can distinguish the seizures from IAE and from ANE. This could allow for the early management of children with influenza in order to prevent morbidity and mortality. The biochemical/hematologic markers lacked specificity.

## Introduction

Influenza in children is a major cause of morbidity and mortality worldwide <sup>1-3</sup>. Annual epidemics in adults and children are associated with an estimated 3-5 million cases of severe illness, and about 290,000-650,000 deaths <sup>2</sup>. Influenza infection is seasonal in temperate countries, with peaks during the winter months, but it has sustained activity throughout the year in tropical climates <sup>1-3</sup>. Infection can be caused by subtypes A, B, and C; influenza A and B are the dominant circulating viruses <sup>3-5</sup>. Minor genetic variations (antigenic drift) are the cause of seasonal variation, and major large-scale reassortments generate novel strains (antigenic shift) with little or no pre-existing immunity in the human population, leading to pandemic strains <sup>1,3,5</sup>. Transmission occurs via respiratory droplets and fomites. The incubation period is 1-4 days<sup>1,3</sup>.

Populations with increased risk for complicated or severe disease course include all children aged 6-59 months, children who have chronic pulmonary or cardiovascular, renal, hepatic, neurologic, hematologic, or metabolic disorders, children with immunosuppression due to medications or disease, children and adoles-

cents who are receiving long-term aspirin therapy, and children with morbid obesity<sup>4</sup>. The major complications include acute otitis media, primary viral pneumonia, and influenza-associated bacterial pneumonia, myositis/rhabdomyolysis, myocarditis, pericarditis, and central nervous system diseases. Among the neurological complications, seizures, acute influenza virus-associated encephalitis (IAE), and acute necrotizing encephalopathy (ANE) indicate critical influenza infection and are associated with significant morbidity and mortality. Among the patients with IAE, 81.8% are 1-5 years of age<sup>6</sup>. IAE presents as convulsions, acute disturbance of consciousness, and coma<sup>7-9</sup>. ANE typically occurs in children <5 years of age and is characterized by rapid progression to encephalopathy, coma, or death within 1-2 days from onset<sup>7,8,10</sup>. The mortality rate of ANE is about 70%<sup>11</sup>.

Because these neurological complications can be managed to some extent, and prognosis can be improved if early treatment is undertaken<sup>11</sup>, predicting the occurrence of neurological complications if influenza is of clinical importance. Some previous studies suggested laboratory indicators, signs, and symptoms that are associated with ANE<sup>12-15</sup>. Yamamoto et al.<sup>16</sup> suggested a severity score for ANE that is able to identify the patients at high-risk of ANE but did not include seizures and IAE.

Therefore, the aim of the present study was to establish and validate an early prediction model to discriminate among neurological complications such as seizures, IAE, and ANE in children with influenza. Such a system could allow identifying the children at high risk of neurological complications and in whom early treatment should be performed in order to improve prognosis.

## Methods

### Study design and patients

This was a retrospective single-center case-control study conducted at Guangzhou Women and Children Medical Center, China, from November 2012 to January 2020. This study was approved by the Ethics Committee of Guangzhou Women and Children Medical Center. All patients signed an informed consent form upon admission.

The inclusion criteria were: 1) children (<18 years); 2) admitted to the hospital with influenza virus infection; and 3) neurological manifestations during hospitalization. The exclusion criteria were: 1) admission >7 days after onset; 2) co-infected with other pathogens; 3) comorbidities like brain trauma, sequelae of viral encephalitis, or metabolic diseases; 4) missing data >30%; or 5) neurological complications other than seizures, IAE, or ANE.

Influenza with seizures was defined as convulsive seizures during fever, consciousness after the seizures, a maximum of two seizure events, and no abnormalities in the cerebrospinal fluid examination and head imaging. IAE was defined as convulsions, acute cognitive impairment, acute disturbance of consciousness, and coma<sup>7-9</sup>, without specific biochemistry and imaging changes<sup>11</sup>. ANE was defined as acute fever, frequent convulsions, acute disturbance of consciousness, coma, and multiple organ failure, with a risk of death<sup>7,8,10</sup>; biochemistry changes are not specific<sup>11</sup>, but imaging shows brain edema and necrosis of thalamus and other deep brain structures<sup>11,17</sup>.

### Observation indexes

Detailed demographic, clinical characteristics at admission, and biochemistry and hematologic indicators of the included patients were extracted from the structured electronic medical records system (EMRS). The earliest value of hematologic indicators within 48 h after admission was used.

### Prediction model development

The prediction model was developed based on random forests (RF), which is an ensemble of decision trees<sup>18-20</sup>. RF is good at describing the relationship between independent and dependent variables with high flexibility and sufficient accuracy<sup>18</sup>. The two main parameters in RF are mtry (the number of random variables used in each tree) and ntree (the number of trees used in the forest). In this model, the mtry

value was the square root of the number of predictors, and the ntree value was 500. The missing values were replaced by the median of each group.

The study patients were split into two separate data sets using 5-fold cross-validation by the RF method: 80% of them were in the training set (the algorithm creation group), and the remaining 20% were in the validation set to obtain unbiased estimates of correct classification rates and variable importance. The equation of the correct classification rate was:

$$\text{Correct classification rate} = \frac{\text{True Seizures} + \text{True IAE} + \text{True ANE}}{\text{Number of patients in the data set}}$$

## Statistical analysis

Categorical variables were presented as counts and percentages, and the differences were analyzed using the chi-square test. Continuous variables were tested for normal distribution using the Shapiro-Wilk test. The continuous variables were presented as medians and interquartile ranges (IQR) based on non-normal distribution, and the differences between groups were analyzed using the Kruskal-Wallis test. All probability values were 2-sided when applicable, and P-values <0.05 were considered statistically significant. Analyses were performed using SAS 9.4 for Windows (SAS Institute, Inc., Cary, NC, USA) and R software (version 3.2.5).

## Results

### Characteristics of the patients

Figure 1 presents the enrollment flowchart. There were 433 patients included in the analysis. Of the 433 patients (294 males, 139 females; median age 2.8 (1.7,4.8) years) with neurological manifestations that occurred during influenza infection, 278 (64.2%) were ultimately diagnosed as seizure, 106 (24.5%) as IAE, and 49 (11.3%) as ANE. The 78 variables, including demographic characteristics, clinical symptoms, and biochemical and hematologic indicators collected for each patient are shown in Table 1. The incidences of in-hospital death of the three groups were 0.4% (1/278; this patient had a chromosomal abnormality), 0% (0/106), and 32.7% (16/49) in the seizure, IAE, and ANE groups, respectively.

### Variable selection

Variable selection was carried out using the different feature subsets RF method. The top 15 variables selected by importance are shown in Figure 2. The larger the importance number is, the more important the variable is. Figure 3 shows the relationship between the cross-validation error and the number of variables. When the number of variables increases to 10, the error achieves a minimum of 0.16. With the number of variables increasing gradually to 78, the error increases. Thus, the final model included 10 features for IAE and ANE prediction: convulsions, procalcitonin (PCT), urea,  $\gamma$ -glutamyl transferase ( $\gamma$ -GT), aspartate aminotransferase (AST), albumin/globulin ratio (A/G),  $\alpha$ -hydroxybutyric dehydrogenase (HBD), alanine aminotransferase (ALT), alkaline phosphatase (ALP), and C-reactive protein (CRP).

### Model development

The influences of the selected variables calculated by the random forest on the seizure are shown in Figure 4. When the number of convulsions was only 1-2 at admission, confidence could be high that the child will only have seizures, but the number of convulsions in children with IAE or ANE could be [?]3 or 0. With increases in PCT, urea,  $\gamma$ -GT,  $\alpha$ -HBDH, ALT, and AST levels, the children were less likely to have only seizures. With the decreases of A/G and ALP, the children were less likely to have only seizures.

### Model validation

The prediction accuracy of the model was internally evaluated by 5-fold cross-validation. The basic information comparing the train and test sets is shown in Table 2, and all the differences in the characteristics of the two data sets were not significantly different.

The prediction model gave a prediction accuracy of 84.2%. In order to examine the performance of the newly developed model, the training model was tested based on the validation set containing 85 patients. The external validation achieved 88.2% accuracy. Based on Table 3, it was less likely that seizure was wrongly classified, but IAE (22.7%, 5/22) was prone to be misdiagnosed as seizure, and a small proportion (4.5%, 1/22) of them was prone to be misdiagnosed as ANE. Of the children with ANE, 22.2% (2/9) were misdiagnosed as IAE, and none were misdiagnosed as seizures.

## Discussion

Neurological complications of influenza are associated with high morbidity and mortality in children<sup>11</sup>. The prognosis could be improved if early treatments are undertaken<sup>11</sup>. Therefore, this study aimed to establish and validate an early prediction model to discriminate among neurological complications such as seizures, IAE, and ANE in children with influenza. The results suggest that this model can distinguish the seizures from IAE and from ANE on patients hospitalized within 7 days of onset. This could allow for the early management of children with influenza in order to prevent morbidity and mortality. The biochemical/hematologic markers lacked specificity.

In this study, the first measurements of biochemical and hematologic indicators within 48 h after admission were evaluated using the RF method, avoiding the problem of model overfitting caused by correlations among the variables. The model has a good ability to distinguish seizures from IAE and ANE early after admission, but for children with IAE, the early hematologic indicators might be misleading, suggesting that the early blood biochemical indicators lack specificity for IAE. In addition, for children with ANE, which has a high mortality rate, attention should be paid to the number of convulsion events before admission and then the changes in biochemical and hematologic indicators.

These results are globally supported by previous studies and by the natural history of IAE and ANE. Indeed, ANE is characterized by frequent convulsions<sup>8,10,21</sup>, as observed in the present study. A previous study showed that a combination of age <4 years, repeated seizures, altered consciousness, and positive Babinski's sign were the high-risk factors for ANE<sup>21</sup>. On the other hand, the literature suggests that there is no specific laboratory marker for ANE<sup>9,22,23</sup>, but elevated AST, elevated glucose, hematuria, proteinuria, and RANBP2 mutations could be associated with ANE<sup>8</sup>. Elevated serum transaminases could also be associated with ANE<sup>9</sup>. In the present study, the likelihood of seizures decreased with the increasing levels of AST, and many studies showed that AST levels increase when soft-tissue necrosis occurs<sup>24-26</sup>. Early research showed that brain dehydration was maximal 30 min after urea injection and resulted in an improvement of cerebral circulation<sup>27</sup>; the increase in urea might be related to the reactive regulation of early cerebral edema. The endothelial cells of the capillaries of the cerebral cortex in rats showed high  $\gamma$ -GT activity<sup>28</sup>, suggesting the possibility of cerebrovascular involvement in early ANE. Activities of  $\alpha$ -HBDH were measured in rats after intermittent exposure to aerogenic hypoxia but no effects on adults<sup>29</sup> and were also associated with edema, ischemic and hemorrhagic changes<sup>30</sup>. In the early stages of influenza, increases in these factors could be a high risk for ANE.

In the present study, the imaging parameters could not be included in the analyses, mainly because of the too wide variety of examination types and examination protocols performed. Future studies about the refinement of the present model should examine the possibility of including imaging variables. Indeed, most cases of ANE display brain edema early in the course of the disease<sup>17</sup>, before the appearance of thalamus necrosis, which is usually at about 3 days after onset<sup>31,32</sup>. The presence of brain imaging features is usually associated with poor prognosis<sup>12-15</sup>. Furthermore, cerebrospinal fluid examination in patients with ANE usually reveals increased amounts of proteins<sup>9</sup>, but the cerebrospinal fluid examination was not performed in all patients. A predictive model for ANE severity included imaging and cerebrospinal features such as brain stem lesions and cerebrospinal proteins<sup>16</sup>. In the present study, older children were more prone to ANE, but age ranked 12th in importance, and only the first 10 variables were included, based on the cross-validation error minimum principle.

Regarding IAE, the accuracy was lower than for ANE. This could be because IAE is, based on the symptoms,

an intermediate condition between seizures and ANE<sup>11</sup>. Indeed, as for ANE, AIE is characterized by convulsions, disturbed consciousness, and coma<sup>7-9</sup>, but it does not progress to death<sup>11</sup>, and no specific biochemical, hematological, and imaging markers can be identified<sup>11</sup>. Nevertheless, vascular injuries in the brain could be found in cadaveric studies<sup>33,34</sup>, but without progression to brain necrosis, as seen in ANE.

This study has limitations. The correct classification rate of seizure is high in our study, but the ability to early diagnosis IAE and ANE still needs to be improved. This could be improved by including imaging characteristics. In addition, for the biochemical and hematological indicators, only the value in the first 48 h was considered, and the eventual changes in those indicators were not considered. Since ANE is a rapidly progressing condition, the exact timing of the evaluations can affect the results. In addition, the study was performed at a single hospital, and the sample size was therefore limited.

## Conclusion

In conclusion, this RF model that was developed in the present study can distinguish seizures from IAE and ANE, with high accuracy. This tool could allow the early and correct discrimination between seizures and ANE, allowing the initiation of early treatments in the identified patients. Nevertheless, biochemical and hematologic characteristics lacked specificity. Additional study is still necessary to refine this model.

## References

- Centers for Disease Control and Prevention. Influenza (Flu). <https://www-cdc-gov.acces.bibl.ulaval.ca/flu/index.htm> . Accessed 05-01-2020. 2020.
- World Health Organization. Influenza. <https://www.who.int/influenza/en/> . Accessed 05-01-2020. 2020.
- Treanor JJ. Influenza (Including avian influenza and swine influenza). In: Bennett J, Dolin R, Blaser M, eds. *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases, eighth edition*. New York: Saunders; 2015:2000-2024.
- Grohskopf LA, Alyanak E, Broder KR, Walter EB, Fry AM, Jernigan DB. Prevention and Control of Seasonal Influenza with Vaccines: Recommendations of the Advisory Committee on Immunization Practices - United States, 2019-20 Influenza Season. *MMWR Recommendations and reports : Morbidity and mortality weekly report Recommendations and reports*. 2019;68(3):1-21.
- Kash JC, Taubenberger JK. The role of viral, host, and secondary bacterial factors in influenza pathogenesis. *The American journal of pathology*. 2015;185(6):1528-1536.
- Wong KK, Jain S, Blanton L, et al. Influenza-associated pediatric deaths in the United States, 2004-2012. *Pediatrics*. 2013;132(5):796-804.
- Togashi T, Matsuzono Y, Narita M, Morishima T. Influenza-associated acute encephalopathy in Japanese children in 1994-2002. *Virus Res*. 2004;103(1-2):75-78.
- Howard A, Uyeki TM, Fergie J. Influenza-Associated Acute Necrotizing Encephalopathy in Siblings. *Journal of the Pediatric Infectious Diseases Society*. 2018;7(3):e172-e177.
- Mizuguchi M, Abe J, Mikkaichi K, et al. Acute necrotising encephalopathy of childhood: a new syndrome presenting with multifocal, symmetric brain lesions. *Journal of neurology, neurosurgery, and psychiatry*. 1995;58(5):555-561.
- Weitkamp JH, Spring MD, Brogan T, Moses H, Bloch KC, Wright PF. Influenza A virus-associated acute necrotizing encephalopathy in the United States. *The Pediatric infectious disease journal*. 2004;23(3):259-263.
- Chen Q, Li P, Li S, Xiao W, Yang S, Lu H. Brain Complications with Influenza Infection in Children. *J Behav Brain Sci*. 2020;10.

12. Mizuguchi M. Acute necrotizing encephalopathy of childhood: a novel form of acute encephalopathy prevalent in Japan and Taiwan. *Brain & development*.1997;19(2):81-92.
13. Okumura A, Mizuguchi M, Kidokoro H, et al. Outcome of acute necrotizing encephalopathy in relation to treatment with corticosteroids and gammaglobulin. *Brain & development*. 2009;31(3):221-227.
14. Wong AM, Simon EM, Zimmerman RA, Wang HS, Toh CH, Ng SH. Acute necrotizing encephalopathy of childhood: correlation of MR findings and clinical outcome. *AJNR American journal of neuroradiology*. 2006;27(9):1919-1923.
15. Kim JH, Kim IO, Lim MK, et al. Acute necrotizing encephalopathy in Korean infants and children: imaging findings and diverse clinical outcome. *Korean journal of radiology*. 2004;5(3):171-177.
16. Yamamoto H, Okumura A, Natsume J, Kojima S, Mizuguchi M. A severity score for acute necrotizing encephalopathy. *Brain & development*. 2015;37(3):322-327.
17. Kirat N, De Cauwer H, Ceulemans B, Vanneste D, Rossi A. Influenza-associated encephalopathy with extensive reversible restricted diffusion within the white matter. *Acta Neurol Belg*. 2018;118(4):553-555.
18. Boulesteix AL, Janitz S, Kruppa J, Konig IR. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *WIREs Data Mining Knowledge Discovery*. 2012;2(6):493-507.
19. Casanova R, Saldana S, Chew EY, Danis RP, Greven CM, Ambrosius WT. Application of random forests methods to diabetic retinopathy classification analyses. *PloS one*.2014;9(6):e98587.
20. Sarica A, Cerasa A, Quattrone A. Random Forest Algorithm for the Classification of Neuroimaging Data in Alzheimer's Disease: A Systematic Review. *Frontiers in aging neuroscience*. 2017;9:329.
21. Centers for Disease C, Prevention. Estimates of deaths associated with seasonal influenza — United States, 1976-2007. *MMWR Morbidity and mortality weekly report*. 2010;59(33):1057-1062.
22. Shah S, Keil A, Gara K, Nagarajan L. Neurologic complications of influenza. *Journal of child neurology*. 2014;29(9):NP49-53.
23. Kondrich J, Rosenthal M. Influenza in children. *Current opinion in pediatrics*.2017;29(3):297-302.
24. Cakir M, Tekin A, Kucukkartallar T, Vatansev H, Kartal A. Enzymatic debridement in necrotizing pancreatitis. *Int Surg*. 2015;100(5):897-902.
25. Huang KC, Hsieh PH, Huang KC, Tsai YH. Vibrio necrotizing soft-tissue infection of the upper extremity: factors predictive of amputation and death. *J Infect*.2008;57(4):290-297.
26. Hacimustafaoglu M, Celebi S, Sarimehmet H, Gurpinar A, Ercan I. Necrotizing pneumonia in children. *Acta Paediatr*. 2004;93(9):1172-1177.
27. Bralet J, Beley P, Bralet AM, Beley A. Comparison of the effects of hypertonic glycerol and urea on brain edema, energy metabolism and blood flow following cerebral microembolism in the rat. Deleterious effect of glycerol treatment. *Stroke*. 1983;14(4):597-604.
28. Budi Santoso AW, Bar T. Postnatal development of gamma-GT activity in rat brain microvessels corresponds to capillary growth and differentiation. *Int J Dev Neurosci*.1986;4(6):503-511.
29. Trojan S, Langmeier M, Maresova D, Pokorny J. Dehydrogenase activity in the blood and brain after adaptation to intermittent hypoxia. *Sb Lek*. 2000;101(1):11-16.
30. Zukiel R, Nowak S, Jankowski R, Moczko J, Michalak S. [Enzymatic activity in cerebrospinal fluid in the monitoring of the brain lesions following intracranial tumors]. *Neurol Neurochir Pol*. 1998;32(2):359-365.
31. Mizuguchi M, Hayashi M, Nakano I, et al. Concentric structure of thalamic lesions in acute necrotizing encephalopathy. *Neuroradiology*. 2002;44(6):489-493.

32. Ormitti F, Ventura E, Summa A, Picetti E, Crisi G. Acute necrotizing encephalopathy in a child during the 2009 influenza A(H1N1) pandemic: MR imaging in diagnosis and follow-up. *AJNR American journal of neuroradiology*.2010;31(3):396-400.
33. Sato T, Moriuchi H. [Influenza-associated encephalopathy]. *Nihon rinsho Japanese journal of clinical medicine*. 2010;68(9):1661-1665.
34. Shiomi M. [Pathogenesis of acute encephalitis and acute encephalopathy]. *Nihon rinsho Japanese journal of clinical medicine*. 2011;69(3):399-408.

## Figure legends

**Figure 1.** Enrollment flow chart. IAE: influenza virus-associated encephalitis; ANE: acute necrotizing encephalopathy.

**Figure 2.** Relationship between the cross-validation error and the number of variables.

**Figure 3.** Importance of the top 15 variables selected by random forests (RF) method. Procalcitonin, PCT;  $\gamma$ -glutamyltransferase,  $\gamma$ -GT; aspartate aminotransferase, AST;  $\alpha$ -hydroxybutyric dehydrogenase,  $\alpha$ -HBDH; alanine aminotransferase, ALT; alkaline phosphatase, ALP; c-reactive protein, CRP; lactate dehydrogenase, LDH; oxygen partial pressure, OPP; prothrombin time, PT.

**Figure 4.** Influence of the selected variables calculated by the random forest method on the seizure.

PCT: procalcitonin;  $\gamma$ -GT:  $\gamma$ -glutamyltransferase; AST: aspartate aminotransferase; A/G: albumin/globulin ratio;  $\alpha$ -HBDH:  $\alpha$ -hydroxybutyric dehydrogenase; ALT: alanine aminotransferase; ALP: alkaline phosphatase; CRP: C-reactive protein.

## Hosted file

Table 1.docx available at <https://authorea.com/users/351892/articles/476329-an-early-prediction-model-to-identify-neurological-complications-of-childhood-influenza-a-random-forest-model>

## Hosted file

Table 2.docx available at <https://authorea.com/users/351892/articles/476329-an-early-prediction-model-to-identify-neurological-complications-of-childhood-influenza-a-random-forest-model>

## Hosted file

Table 3.docx available at <https://authorea.com/users/351892/articles/476329-an-early-prediction-model-to-identify-neurological-complications-of-childhood-influenza-a-random-forest-model>







