EPIDEMIOLOGY AND CLINICAL FEATURES OF SARS-COV-2 AND HERPESVIRUS INFECTION DURING CHILDHOOD

Ph.D. THESIS

Andrea Tímea Takács MD

Supervisor

Gabriella Terhes MSc, PhD

Department of Pediatrics and Pediatric Health Center

University of Szeged

Szeged

2023

LIST OF FULL PAPERS RELATED TO THE SUBJECT OF THE THESIS

I.   **Takács A,** Bukva M, Bereczki Cs, Burián K, Terhes G. Diagnosis of Epstein-Barr and cytomegalovirus infections using decision trees: an effective way to avoid antibiotic overuse in paediatric tonsillopharyngitis.
     BMC Pediatr. 2023 Jun;23(1):301.
     IF: 2.567 (Pediatrics, Perinatology and Child Health Q1)

II. **Takács AT,** Bukva M, Gavallér G, Kapus K, Rózsa M, Bán-Gagyi B, Sinkó M, Szűcs D, Terhes G, Bereczki C. Epidemiology and clinical features of SARS-CoV-2 infection in hospitalized children across four waves in Hungary: A retrospective, comparative study from March 2020 to December 2021.
     Health Sci Rep. 2022 Nov;5(6):e937.
     IF: - (Medicine Q2)

III. **Takács A**, Szűcs D, Terhes G. [Exudative tonsillitis in children]. How can we reduce the unnecessary antibiotic consumption?]
     Orv. Hetil. 2020 Jan;161(2):50-55. *(Hungarian)*
     IF: 0.54 (Medicine Q4)

1. INTRODUCTION

The global healthcare landscape is navigating through two significant but distinct challenges: the alarming escalation of antimicrobial resistance (AMR) and the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic. While each crisis poses challenges and implications for public health, healthcare infrastructure, and global economies, they intersect in complex ways. The SARS-CoV-2 pandemic and cytomegalovirus- (CMV) or Epstein-Barr virus (EBV)- induced upper respiratory tract infections have an impact on antibiotic usage, which in turn has implications for the already critical issue of AMR.

1.1   The growing threat of antibiotic resistance

The spread of multidrug-resistant pathogens resulting from the unwarranted use of antibiotics is a worldwide problem. AMR remains one of the top 10 global public health threats facing humanity. In 2019 alone, AMR was associated with the deaths of nearly 5 million people, with methicillin-resistant *Staphylococcus aureus* accounting for over 100,000 of these fatalities. Projections suggest that AMR could claim up to 10 million lives

annually by 2050. The distribution of resistant bacteria varies considerably in Europe, but the Southern and Eastern regions are the most affected.

To effectively mitigate the risk of AMR, a multifaceted approach is essential. One primary driver of AMR is the inappropriate use of antibiotics in treating viral infections. For instance, despite most of the childhood tonsillitis being of viral origin, antibiotics are frequently prescribed in contravention of international guidelines. This not only constitutes an ineffective treatment strategy but also significantly contributes to the escalating problem of AMR. Acute pharyngitis, commonly associated with viral infections like adenoviruses, influenza, CMV and EBV often receives unnecessary antibiotic treatment out of an unfounded fear of bacterial complications. Bacteria, such as group A streptococci (GAS), cause only 15-30% of cases.

This antibiotic abuse also characterises infectious mononucleosis (IM), which CMV or EBV causes. Although IM has several distinguishing signs (exudative pharyngitis, hepatosplenomegaly, and lymphadenopathy), other pathogens, such as bacterial infections, are difficult to identify from IM. In practice, 53.10-72.69% of IM patients are given antibiotics for a bacterial infection.

Based on the preceding, it is becoming increasingly crucial to distinguish between viral and bacterial pharyngitis or to utilise algorithms that can accurately identify particular viral/bacterial illnesses.

1.2 The SARS-CoV-2 pandemic and antibiotic resistance

The SARS-CoV-2 pandemic, as of September 2023, has resulted in over 770 million confirmed cases and nearly 7 million deaths. As reported by the Centers for Disease Control and Prevention (CDC), between March and October 2020, almost 80% of hospitalised patients with Coronavirus disease-19 (COVID-19) received antibiotics. About half of them have received ceftriaxone, usually prescribed in combination with azithromycin. In the first year of the pandemic, over 29,400 individuals died due to antimicrobial-resistant bacteria. Nearly 40% of these people got the infection in the hospital. Children and adolescents have generally experienced less severe forms of SARS-CoV-2 infections, with a lower mortality rate compared to adults. According to the Hungarian database, 1 256 415 cases of COVID-19 have been confirmed in Hungary by the end of December 2021. In contrast to the high adult mortality rate (n = 39 186), only 14 children succumbed to the disease in Hungary by December 2021. Although the several waves of the SARS-CoV-2 pandemic have been extensively studied, few studies have examined the symptoms and consequences of the

succeeding four waves, particularly in hospitalised children of various ages. However, this would be crucial to reduce the spread of AMR in several ways: children infected with SARS-CoV-2 often present with symptoms different from adults, making diagnosis difficult. Children with severe COVID-19 symptoms may require hospitalisation, where the risk of acquiring nosocomial infections, including infections caused by antibiotic-resistant bacteria, is higher. Understanding how COVID-19 affects children can help develop targeted infection control measures, reducing the need for broad-spectrum antibiotics that drive resistance.

1.3 Focus of the thesis

In the light of the above, the focus of the thesis is divided into two aspects, which are intended to address the shortcomings outlined above.

- This paper proposes a machine learning decision tree that can identify tonsillopharyngitis caused by EBV/CMV and thus could reduce unnecessary antibiotic use for these two pathogens.

- This dissertation would bridge the current knowledge gap in paediatrics on the epidemiology of COVID-19. Research on the epidemiology and clinical course of the virus could help better understand its characteristics, which could contribute to the fight against AMR in the longer term.

2. AIMS

2.1 Assessment of EBV and CMV-triggered upper respiratory tract infections during childhood

- Retrospective evaluation of laboratory parameters and clinical signs to distinguish these patients from infections caused by other pathogens

- To determine whether the introduction of the decision tree method could reduce the use of antibiotics

2.2 Retrospective assessment of a population of children hospitalised for COVID-19 disease

- To examine the age distribution of patients in each wave

- Assessment of the presence of clinical signs and symptoms in childhood and their age distribution

- To identify the distribution of typical clinical signs and symptoms by wave

- Evaluation of the characteristics of immunosuppressed patients requiring hospitalisation
- To examine the factors that affect the length of hospitalisation

3. MATERIALS AND METHODS

3.1 Materials and methods on the survey of diagnosis of Epstein-Barr and cytomegalovirus infections using decision trees in paediatric tonsillopharyngitis

3.1.1 Patients

This research was executed between January 2016 and December 2017. Patients with tonsillopharyngitis with EBV and CMV serologic tests were included (n = 242). They were admitted to the Department of Paediatrics at the University of Szeged (SZTE) in Hungary for outpatient and inpatient care. Past medical history, clinical findings on presentation, indication for EBV and CMV serologies, biochemical test results (complete blood count, liver function tests, and C-reactive protein), treatments, and demographic data were all analysed. At the time of their first visit, all the patients were between the ages of 0 and 18. We identified upper respiratory tract infection: the patient had a fever, sore throat, cervical lymphadenopathy, and inflamed pharynx with or without exudate. Patients were split into two groups (EBV/CMV and non-EBV/CMV) based on whether they confirmed EBV or CMV.

3.1.2 Microbiological investigations

To detect CMV and EBV-specific antibodies enzyme-linked immunosorbent assay (ELISA) were utilised. Anti-Viral capsid antigen (VCA) Immunoglobulin-M (IgM) positive/borderline results and anti-VCA (Immunoglobulin-G) IgG positive results with negative anti- Epstein-Barr nuclear antigen (EBNA) IgG test results confirm primary EBV infection. Anti-CMV IgM was positive, and seroconversion was detected in the second specimen of the primary CMV infection. Matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS) was used to identify the isolated strain in the presence of *Streptococcus pyogenes*.

3.1.3 Biochemical tests

On a Sysmex XE-2100 automated analyser we evaluated the whole blood count using the Sysmex reagent. The International Federation of Clinical Chemistry (IFCC)

standardised technique for alanine aminotransferase (GPT), aspartate aminotransferase (GOT), and C-reactive protein (CRP) analysis was used on a ROCHE Modular P800 analyser (Roche, Switzerland).

3.1.4 Statistical analysis

Symptoms were applied as categorical variables in the classification trees, and laboratory findings were used as continuous variables. 80% of the data was utilised as a training set, while the remaining 20% was used to test the built trees. Three classification trees were created: one for symptoms, one for blood test findings, and one for all these data. All models were given an age. We employed Welch's and Fisher's exact tests for continuous and discrete variables for univariable statistical analyses. For the difference in means and the odds ratio (OR), we estimated 95% confidence intervals (CI). The GPT (U/L) variable was subjected to receiver operating characteristic (ROC) and the area under the curve (AUC) analysis and associated 95% confidence intervals were obtained. The excellent Youden index (sensitivity + specificity - 100) indicated the best cut-off point. The statistical analysis, machine learning algorithms, and visualisations were created using Orange 3.32 and GraphPad Prism 9. A value of $p < 0.05$ was considered significant. The local ethics commission accepted the investigation (ethical permission number: 124/2016 SZTE).

3.2 Materials and methods on the survey of epidemiology and clinical features of SARS-CoV-2 infection in hospitalised children across four waves from March 2020 to December 2021

3.2.1 Patients

This survey was conducted retrospectively between March 2020 and December 2021. The four COVID-19 waves we evaluated in our study were as follows, based on the occurrence of SARS-CoV-2 mutations in Hungary. We investigated data from all patients who required hospitalisation owing to positive SARS-CoV-2 nucleic acid amplification test (NAAT) or rapid antigen test findings at the University of Szeged in Hungary. The study included patients who needed hospital treatment for other reasons (such as surgical management, poisoned children, trauma, and so on) and tested positive for coronavirus. Using medical records, data analysis included demographic data, past medical history, clinical findings at admission, length of hospitalisation, and complications. All patients were between 0 and 18 at the time of admission. The COVID-19 case definition from World Health Organisation (WHO) was applied.

3.2.2 Microbiological investigations

For SARS-CoV-2 nucleic acid detection, nasopharyngeal and throat swabs were collected using Universal Transport Medium (UTM-RT). For ribonucleic acid (RNA) isolation, the QIAamp 96 Virus QIAcube HT Kit was used. The evaluation of real-time polymerase chain reaction (PCR) findings was performed by the manufacturer's instructions. We used the VivaDiag™ Pro SARS-CoV-2 Ag Rapid Test and the Panbio™ COVID-19 Ag Rapid Test Device to detect the nucleocapsid protein from SARS-CoV-2 in nasal or oropharyngeal swab specimens.

3.2.3 Statistical analysis

The relative risk (RR) values and confidence intervals were obtained using the X^2 test with Koopman's asymptotic method. The Cox proportional-hazards model with enter method examined the impacts of symptoms, complications (with indicator contrast), and age groups (with Helmert and reverse Helmert contrast) on hospitalisation time. Cox proportional-hazard models' hazard ratios (HR) with 95% confidence intervals (CI) show increased or reduced hospitalisation time. The Kaplan-Meier method with the log-rank test was used to evaluate the influence of immunocompromised status on hospitalisation time. A heatmap was created to show the incidence of symptoms and consequences across age groups. Principal component analysis (PCA) was performed using the z-scored prevalences. P-values $< .05$ were considered statistically significant. The local ethics committee accepted the investigation (the ethical permission number: 222/2021).

4. RESULTS

4.1 Results of survey of diagnosis of Epstein-Barr and cytomegalovirus infections using decision trees in paediatric tonsillopharyngitis

4.1.1 More than a third of patients with confirmed EBV and/or CMV infection

Infection by EBV and/or CMV was confirmed in 37.60% of patients (91 of 242). The EBV/CMV group was allocated to these individuals. Serology revealed acute EBV in 30.17% of patients (73 of 242) and primary CMV infections in 2.48% (6 of 242) in the EBV/CMV group. Anti-CMV IgM and anti-EBV VCA IgM were positive in 4.96% (12 of 242) of all patients. In the case of EBV/CMV, 40.70% (37 of 91) of patients had a throat swab cultured. Only three patients tested positive for *S. pyogenes*, *Streptococcus agalactiae* or *Candida albicans*. Patients with no confirmed EBV and/or CMV infection (62.40%, 151

of 242) were allocated to the non-EBV/CMV group. Of 151 patients, 60 had throat bacteriological cultures; only five individuals tested positive for *S. pyogenes*, and one tested positive for *Candida albicans*. Adenoviral infection was confirmed in 5 individuals, whereas influenza A and respiratory syncytial virus (RSV) infections were confirmed in 2 and 1 patient, respectively.

4.1.2 Lymphadenopathy, hepatosplenomegaly exudation, and sore throat are characteristic symptoms of EBV/CMV group

As the symptoms were examined, it was shown that the prevalence of lymphadenopathy was 27.02% higher in the EBV/CMV group (81.32% vs 54.30%) than in the non-EBV/CMV group. Hepatosplenomegaly was also 26.24% higher in the EBV/CMV group (56.04% versus 29.80%) than in the non-EBV/CMV group. Exudation was seen 18.26% more commonly in EBV/CMV patients. Cough and nasal discharge (OR = 0.41 [95% CI: 0.22 to 0.77], p = 0.005) were observed to be more prevalent in the non-EBV/CMV group. See Table 1. for all symptom values.

Symptoms	Prevalence in EBV/CMV	Prevalence in non-EBV/CMV	OR	$p*$	95% CI for OR (lower to upper)
Fever	60.00%	66.89%	0.96	0.89	0.55 to 1.66
Lymphadenopathy	81.32%	54.30%	3.66	< 0.0001	1.98 to 6.79
Cough	16.48%	31.79%	0.42	0.009	0.22 to 0.81
Nasal discharge	17.58%	34.44%	0.41	0.005	0.22 to 0.77
Exudation	72.53%	54.30%	2.22	0.006	1.27 to 3.89
Sore throat/dysphagia	63.74%	47.02%	1.98	0.01	1.16 to 3.38
Hepatosplenomegaly	56.04%	29.80%	3	< 0.0001	1.75 to 5.16
Rash	14.29%	15.23%	0.93	0.99	0.44 to 1.94
Abdominal symptom	26.37%	31.79%	0.77	0.39	0.43 to 1.37

*: *p-value* from Fisher's exact test

Table 1. *Prevalence of symptoms among the EBV/CMV and non-EBV/CMV group*

4.1.3 Elevated levels of GOT and GPT were identified in the EBV/CMV group

Examining blood test results and age, it was shown that EBV/CMV patients had higher GOT and GPT levels, as well as the highest average age (Figure 1). GPT levels in the EBV/CMV group were 8.05 times greater than in the non-EBV/CMV group. Similarly, the

average GOT level in EBV/CMV patients was found to be 3.54 times higher, and they were 1.53 times older.
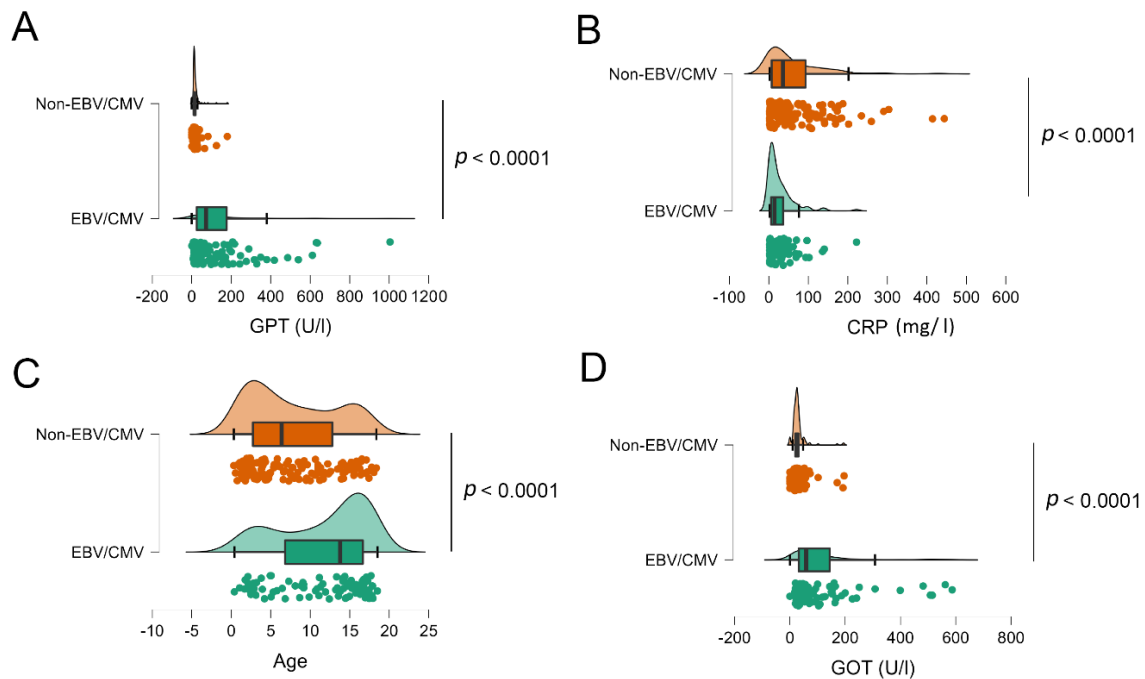

Figure 1. Raincloud plot for continuous variables

The raincloud plot shows the density and quartiles of the GPT (A), CRP (B)(mg/L), age (C), and GOT (U/L) variables. The different colours indicate the EBV/CMV (in green) and non-EBV/CMV (in orange) groups.

4.1.4 Blood test results are more suitable for identifying EBV/CMV infection than symptoms

The data set was divided into training and test sets to generate the classification trees, with 80% of patients allocated to the training set (193 of 242) and 20% assigned to the test set (49 of 242). The train set included 62.70% non-EBV/CMV patients (121 of 193) and 37.30% EBV/CMV patients (72 of 193). The non-EBV/CMV group had 61.27% (31 of 49) and the EBV/CMV group had 36.73% (18 of 49) of the patients in the test set, respectively.

Following data segmentation, three classification trees were constructed based on symptoms, blood test results, and all variables. The algorithm chose sore throat/dysphagia, nasal discharge, exudation, cough, hepatosplenomegaly, and lymphadenopathy to form the symptom-based tree. The best results were obtained with four levels (Figure 2A). Running the model on train dataset, gave an average classification efficiency of 83.80% (Figure 2B). The algorithm accurately categorised 75.00% of EBV/CMV patients and incorrectly classified 25% into the non-EBV/CMV category.

In the test set, the model correctly diagnosed 68.40% of EBV/CMV patients while wrongly identifying 31.60%, for overall classification accuracy of 72.55% (Figure 2C). The positive and negative predictive values calculated for the test data set were 63.16% and 80.00%, respectively.

The co-presentation of hepatosplenomegaly, exudation, and lymphadenopathy were discovered to be the most significant signs in the EBV/CMV group (Figure 2A). This combination was found in 40 patients in the train set, with 80.00% belonging to the EBV/CMV group. The absence of hepatosplenomegaly but the presence of nasal discharge and sore throat/dysphagia, on the other hand, revealed a very distinctive combination for the non-EBV/CMV group. This combination was identified in 34 individuals, 94.10% (32 of 34) of whom were non-EBV/CMV positive.
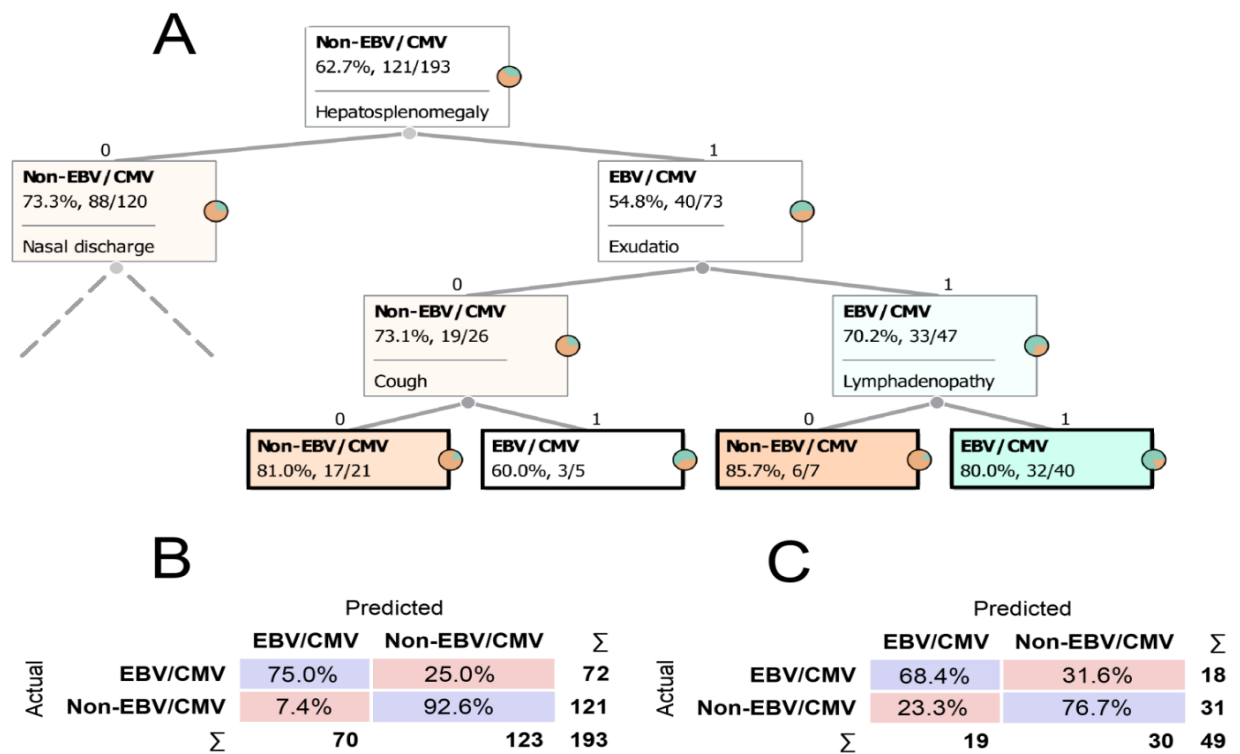
Figure 2. *Classification tree based on symptoms*

Part of the symptom-based classification tree (A): the nodes (rectangles) show the name and proportions of the most frequent group in each step. In the node at the top of the tree, the entire train set (n = 193) is still unpartitioned, in which non-EBV/CVM cases were the most common (62.70%, 121 of 192 A 0 denotes the absence of symptoms, while a 1 denotes their presence. Different intensities of the orange and green colours represent the percentage of non-EBV/CMV and EBV/CMV patients, respectively. Terminal nodes are marked with a black frame. The grey dashed line indicates that only part of the tree is shown. Confusion matrices for test (B) and train set (C): each row represents the instances in an actual class while each column represents the instances in a predicted class. Diagonally, the percentage of the correct classification is shown in blue. The percentage of errors is indicated in red. For example, a total of 72

The algorithm chose the GPT, CRP, GOT, and age to form the classification tree based on the blood test values. Four levels were used to achieve the most significant results (Figure 3A). When run on the train dataset, the model had an average classification accuracy of 85.25% (Figure 3B). The model correctly categorised 82.20% of EBV/CMV patients and incorrectly classified 17.80% of non-EBV/CMV patients. In the test set, the model correctly diagnosed 83.30% of EBV/CMV patients while incorrectly classifying 16.70%, for an overall classification accuracy of 86.80% (Figure 3C).
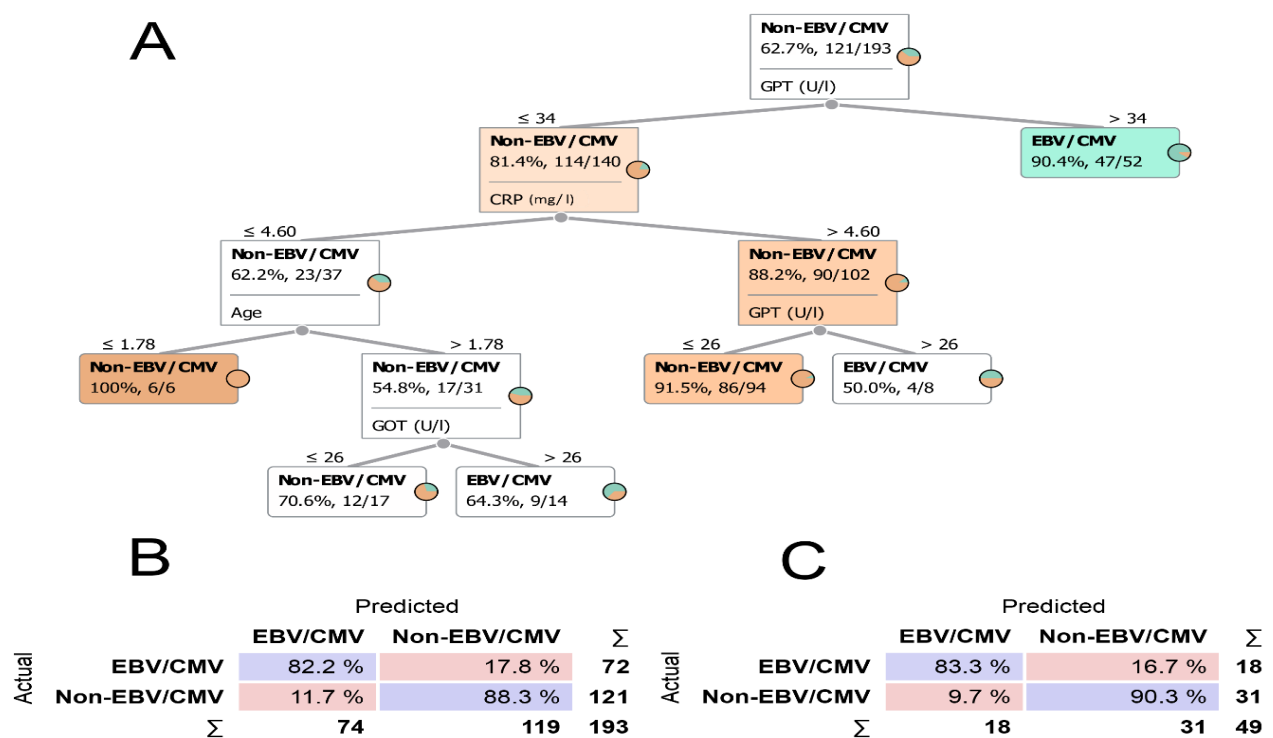
Figure 3. *Classification tree based on the blood test values*

Part of the blood test-based classification tree (A): the nodes (rectangles) show the name and proportions of the most frequent group in each step. In the node at the top of the tree, the entire train set (n = 193) is still unpartitioned, in which non-EBV/CVM cases were the most common (62.70%, 121 of 192). At each step, patients are separated according to the given blood test to form more homogeneous nodes. Different intensities of the orange and green colours represent the percentage of non-EBV/CMV and EBV/CMV patients, respectively. Terminal nodes are marked with a black frame. Confusion matrices for test (B) and train set (C): each row represents the instances in an actual class, while each column represents the instances in a predicted class. Diagonally, the percentage of the correct classification is shown in blue. The percentage of errors is indicated in red. For example, a total of 72 EBV/CMV patients are shown in panel B, and of these 72, 82.2% were classified by the model into the EBV/CMV group and 17.8% into the Non-EBV/CMV group.

The positive and negative predictive values were 83.33% and 90.32%. According to the decision tree, 90.40% (47 of 52) of patients with a GPT level of 34 U/L belong to the EBV/CMV group.The non-EBV/CMV group can be divided into many branches, although 91.5% of patients (86 of 94) belonged to the non-EBV/CMV group in the node with the highest number of patients, with CRP levels over 4.60 but GPT values below 26.00.

4.1.5   The GPT could be a stand-alone marker for EBV/CMV infection

Based on the ROC analysis of the GPT, the AUC value was found to be 0.86 (95% CI: 0.81 to 0.91, $p < 0.0001$), the best cut-off point defined by the highest Youden index was 24.50 U/L, and the corresponding sensitivity and specificity were 88.44% and 75.82%.

4.1.6   Antibiotic treatments can be significantly reduced using our decision tree

Physicians requesting EBV and/or CMV serology considered viral infection in all patients with sore throat. Despite this, antibiotics were provided to 86.8% of EBV/CMV patients (79 of 91). The classification test dataset contained 36.73% EBV/CMV patients (18 of 49). 83.33% of EBV/CMV patients (15 of 18) received antibiotic therapy during clinical care.
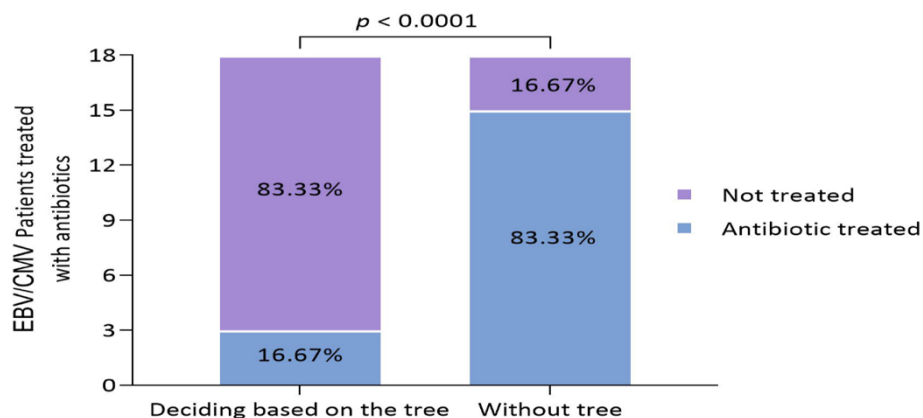

Figure 4. *The proportion of the antibiotic-treated patients in the test data set (n = 18)*

The figure shows the actual proportion of EBV/CMV patients treated with antibiotics in the test data set and the hypothetical proportion that would have been obtained using the classification tree. The p-value is from Fisher's exact test

However, if the blood test-based classification tree had been used to diagnose 13 EBV/CMV patients, they may have avoided antibiotic therapy (Figure 2A). This equates to a 66.66% decrease in the number of EBV/CMV patients receiving antibiotics ($p <$

0.0002) (Figure 4), and none of the patients with positive bacterial cultures would be deprived of antibiotic treatment.

4.2 Results of survey of epidemiology and clinical features of SARS-CoV-2 infection in hospitalised children across four waves from March 2020 to December 2021

4.2.1 Characteristics of COVID-19 patients

Based on medical records, we reviewed data from 358 SARS-CoV-2 infected children who required hospitalisation between March 2020 and December 2021. SARS-CoV-2 infection was confirmed in 55.3% of 358 patients using a rapid antigen test. Females made up 53.35% of the study's participants. More than half (55.6%, 199 of 358) of the patients are under the age of three, and one-third (33.8%, 121 of 353) are under the age of one year. Only one child required in-patient therapy during the first wave (which began in March 2020) (Figure 5B). The second wave (starting in June 2020) was dominated mainly by the Beta variant (B.1.351), with 94 patients requiring in-patient treatment, 13 of whom were not hospitalised owing to the infection. The number of patients in the second wave peaked in December 2020 (n = 34). In the third wave, which began in February 2021 and was caused by the Alpha variant (B.1.1.7), 71 patients were reported, with ten being hospitalised for other reasons. In March 2021, the most patients (n = 35) were registered during the third wave. One hundred ninety-two people were hospitalised during the fourth wave (beginning in July 2021) caused by the Delta variant (B.1.617.2) and 12 of them were admitted for reasons other than infection. The fourth wave's patient population peaked in November 2021 (n = 107) (Figure 5A).

4.2.2 COVID-19 patient with underlying diseases

The comorbidity data of 358 individuals was examined. Seventy-nine children (22.06%) had underlying diseases such as chronic lung disease (n = 35), haematological malignancy (n = 10), neurologic disorders (n = 9), renal comorbidities (n = 7), gastrointestinal abnormalities (n = 4), congenital heart defects (n = 3), severe immunological abnormalities (n = 2), and others (n = 9) (Table 3).
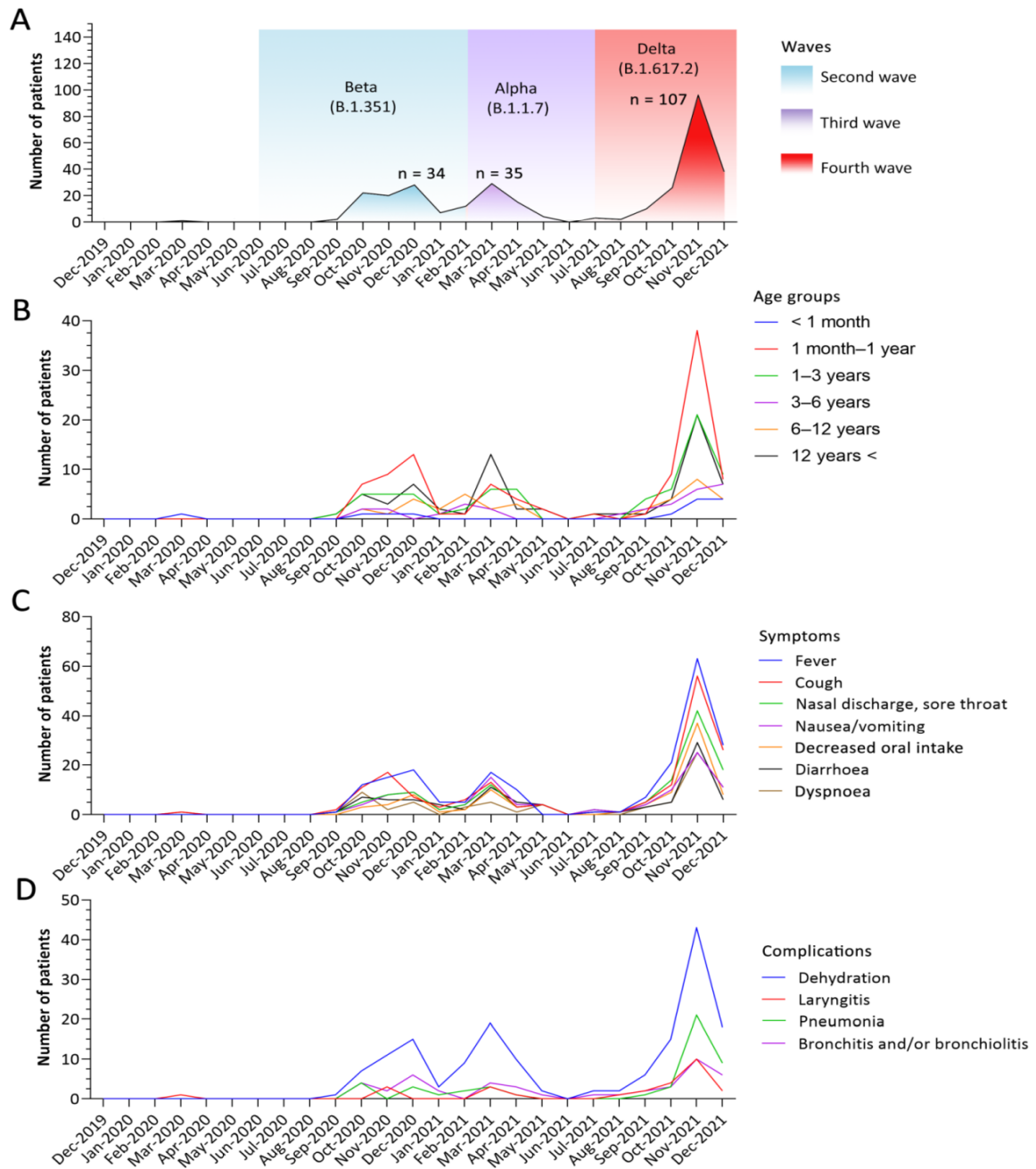
Figure 5. *Time series data*

A)   Patient number distribution over timer (the highest number of patients in a given wave is denoted by n; colours represent the period of each wave based on the Hungarian database (16); the first wave was not present due to the low patient number), B) age distribution of hospitalised patients over time, C) dominating symptoms in COVID-19 patients over time, D) complications in COVID-19 patients over time.

Type of comorbidity	Comorbidity	Number of patients (%[a])	Immunosuppressed (%[b])
Pulmonology (n =35)	Obstructive pulmonary disease	27 (7.5)	
	Bronchopulmonary dysplasia	2 (0.6)	
	Congenital pulmonary abnormalities	2 (0.6)	1 (50)[c]
	Recurrent laryngitis or other laryngeal disorder	4 (1.1)	
Cardiology (n =3)	Congenital heart defect	3 (0.8)	
Haematology (n =10)	Acute lymphoid leukaemia	9 (2.5)	9 (100)
	Immune thrombocytopenia	1 (0.3)	
Immunology (n =2)	Severe combined immunodeficiency	1 (0.3)	1 (100)
	Mannose-binding lectin deficiency	1 (0.3)	1 (100)
Gastrointestinal (n =4)	Inflammatory bowel disease	2 (0.6)	2 (100)
	Gastro-oesophageal reflux	2 (0.6)	
Nephrology (n =7)	Nephrotic syndrome	2 (0.6)	2 (100)
	End-stage renal disease, kidney transplant patients	2 (0.6)	2 (100)
	Congenital kidney disease	3 (0.8)	
Neurology (n =9)	Congenital central nervous system disorders	9 (2.5)	1 (11.11)[c]
Other (n =9)	Obesity	4 (1.1)	
	Hypertony	2 (0.6)	
	Other metabolic disorder	1 (0.3)	
	Comel-Netherton syndrome	1 (0.3)	1 (100)
	Noonan syndrome	1 (0.3)	

Table 3. *The incidence of comorbidity and immunosuppressed patients*
[a]Percentage of all patients (n = 358); [b]percentage of patients with the given comorbidity;
[c]due to other disorders.

4.2.3 Recorded clinical symptoms

Almost 10% of cases (9.7%, 35 of 358) required in-patient treatment for reasons other than COVID-19, including appendicitis (n = 8), traumatic injury (n = 5), diabetic ketoacidosis (n = 3), psychiatric disorder (n = 2), intoxication (n = 2), onco-haematological treatment (n = 2), nephrotic syndrome relapse (n = 1), pyelonephritis (n = 1), urinary tract obstruction (n = 1), oral bleeding (n = 1) and observation for other cases (n = 3). Six asymptomatic individuals needed to be isolated owing to COVID-19. The study comprised 323 COVID-19 patients who were hospitalised due to more severe SARS-CoV-2 infection symptoms (COVID-19-related hospitalisation).

We found that ten symptoms had a prevalence of more than 10%: fever, cough, nasal discharge, nausea/vomiting, reduced oral intake, diarrhoea, dyspnoea, weakness and fatigue, abdominal pain, and headache. In addition, 13 respiratory, systemic, neurological, and other symptoms with a less than 10% prevalence were found. Figure 6A summarises the prevalence of symptoms. The prevalence of the most common symptoms among 18 immunocompromised showed slight differences (fever, 67.4%; cough, 47.3%; nasal discharge, 39.1%; nausea/vomiting, 38.6%; decreased oral intake, 26.2%; diarrhoea, 29.5%; dyspnoea, 27.2%; weakness and fatigue, 22.4%; abdominal pain, 17.3%; headache, 19.7%). However, because the quantity of observations in the two groups varied significantly, we cannot make definite conclusions or acquire trustworthy *p* values.
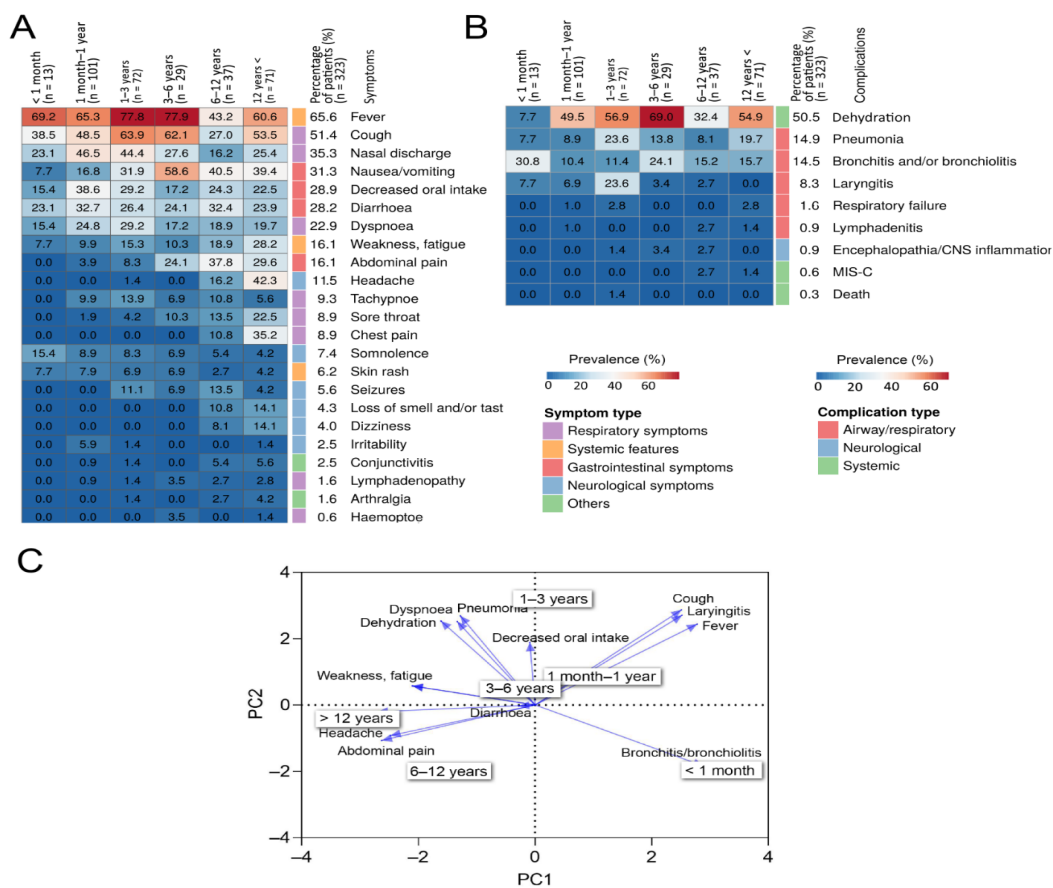

Figure 6. *Characteristics of different age groups*

(A) The heatmap shows the symptoms' prevalence in the investigated age groups. (B) The heatmap shows the prevalence of the complications in the different age groups. (C) The biplot from the principal component analysis reveals the relationship between the symptoms, complications and different age groups. Arrows indicate symptoms and complications. Arrow length and direction are informative: longer arrows are more specific, while shorter arrows indicate more general symptoms and complications. Symptoms and complications close to each other at right angles are not correlated. Angles smaller than right angles indicate a positive correlation, while angles larger than right angles indicate a negative correlation.

4.2.4   Recorded complications

Dehydration, pneumonia, bronchitis, and laryngitis were the four consequences with an incidence of more than 8% among the nine evaluated. More respiratory, neurological, and other symptoms were noted with a reduced frequency. One instance of infection in the 1-3 year age group was fatal during the four waves. Figure 7B summarises the prevalence of symptoms. The prevalence of the most common complications among the 18 immunocompromised showed slight differences (dehydration, 62.4%; pneumonia, 18.3%; bronchitis, 17.7%; laryngitis, 14.7%). However, as the number of observations in the two groups differs remarkably, we cannot draw any firm conclusions and obtain reliable p values.

4.2.5   Characteristic symptoms and complications of the age groups

The prevalence of the 23 symptoms documented varied by age group. All 23 symptoms were seen in youngsters above the age of 12. In comparison, only 43.5% of the symptoms (10 of 23) were observed in newborns. The most complications occurred in children aged 6 to 12, with newborns accounting for 44.4% (4 of 9). The PCA biplot (Figure 7C) revealed a trend in the incidence of symptoms with age. Patients above the age of three reported more frequent symptoms such as headache, stomach discomfort, weakness, and weariness, with a link in prevalence. This might be explained by the fact that the impression of these symptoms is subjective, and older children can better express themselves. According to the PCA biplot, cough, fever, and laryngitis are more common in children under 12. Bronchitis and bronchiolitis have also been identified as frequent problems in children under the age of one month.

4.2.6   Characteristic symptoms and complications of the four waves

The age distribution of patients varies by wave. Children aged over one month and under one year were the most exposed in the second and fourth waves (34%, 32 of 94 and 30.7%, 59 of 192, respectively), although the age distribution was even in the third wave (Figure 5B). When the symptoms were evaluated throughout time, fever was shown to be the most common symptom in each wave (Figure 5C, Figure 7A), and it dominates the clinical scenario. Cough was the second most prevalent symptom in the fall waves of 2020 and 2021, followed by nausea/vomiting in the spring wave of 2021, associated with the Alpha variant (B1.1.7.). Compared to other waves, the fourth wave showed the highest prevalence of symptoms, followed by the third and second waves. According to the PCA biplot, the fourth wave showed distinct symptoms and consequences, such as decreased oral intake, nasal discharge, fever, cough, dyspnoea, pneumonia, and laryngitis (Figure 7C). Cough was 29.3% more prevalent in the fourth

wave compared to the second and third waves combined ($p = 0.022$). At the same time, the prevalence of nasal discharge was 56.2% greater than in the second and third waves combined ($p = .009$). Weakness and fatigue were shown to be common symptoms in the fourth and third waves, with the second wave being 54% lower than the third and fourth waves combined ($p = 0.017$) (Figure 7A, B, C).
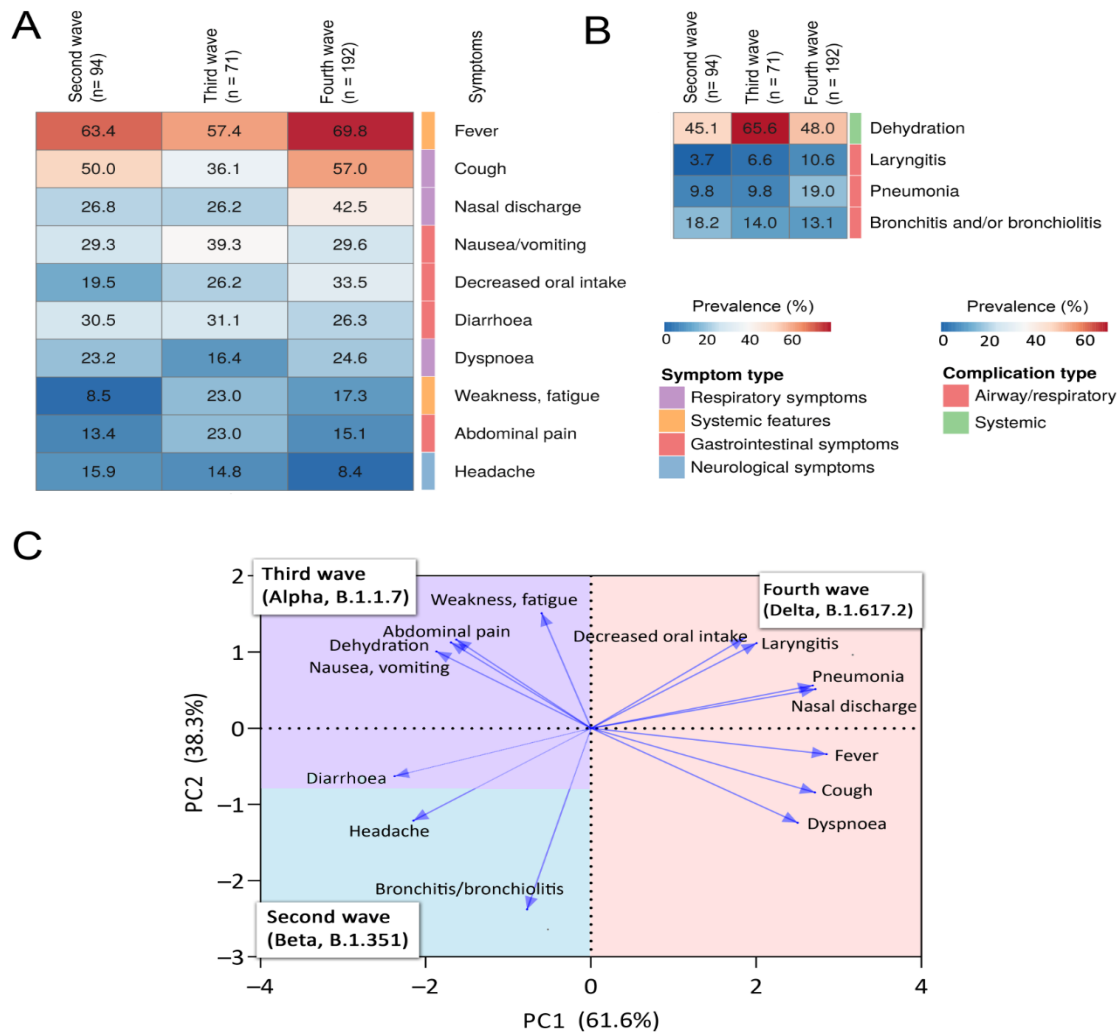

Figure 7. *Characteristics of different waves*

(A) The heatmap shows the prevalence of the symptoms in the different waves. The columns and rows represent the different waves and symptoms, respectively. The heatmap represents the symptoms with a prevalence of over 10%. (B) The heatmap shows the prevalence of the complications in the different waves. The heatmap represents the complications with a prevalence of over 8.3%. (C) The biplot from the principal component analysis reveals the relationship between the symptoms, complications, and the different waves. Arrows indicate symptoms and complications. Arrow length and direction are informative: longer arrows are more specific, while shorter arrows indicate more general symptoms and complications. Symptoms and complications close to each other at right angles are not correlated. Angles smaller than right angles indicate a positive correlation, while angles larger than right angles indicate a negative correlation. The different colours in the plot represent the waves.

The third wave was distinguished by associated stomach discomfort, dehydration, and nausea/vomiting. Dehydration was the most common consequence, with a 40.0% greater frequency than in the second and fourth waves combined ($p = 0.023$) (Figure 7A, B, C).The highest percentages of diarrhoea, headache, bronchitis, and bronchiolitis are reported in the second wave, but RR values are not substantially different from the other waves (Figure 7A, B, C).

Remdesivir antiviral treatment was administered to sixteen children. In the second wave, 2.13% (2 of 94) of patients were affected; in the fourth wave, 7.23% (14 of 192) were involved. This medication was prescribed for 11 patients with pneumonia, 1 with obstructive bronchitis, and 1 with bronchiolitis caused by RSV coinfection. The medication was introduced after one further instance of encephalopathy and one case of MISC. The indication for therapy was not recognised in one patient.

During the four waves, 25 patients received oxygen In the second wave, 5.32% (5 of 94) of patients were affected, 5.63% (4 of 71) of children were involved in the third wave, and 8.33% (16 of 192) were affected in the fourth wave. One patient required mechanical ventilation, another required non-invasive ventilation, and a third required both. Twenty-two people were provided oxygen through a nasal cannula or mask.

4.2.7   Hospitalisation time among children

The entire mean and median length of stay in the hospital were predicted to be and 4 days, respectively. The worst results were reported in the 1-month age group, where more than half of the patients required in-patient treatment for more than five days and 15% for more than seven days. A patient with severe COVID-19 pneumonia was hospitalised for 39 days (37 days in the Intensive Care Unit).

The Cox proportional-hazards model was used to assess the influence of the patient's age group on hospitalisation time, where HR values represent the increase or reduction in hospitalisation time.

The worst HR value was measured in the < 1 month group, in which the patients spend 20% more time in clinical care compared to other age groups (HR = 0.80, 95% CI: 0.49 to 1.40, $p = 0.21$). However, this 20% increase is not statistically significant because the confidence interval for all other groups comprises 0. The Cox proportional-hazard model results revealed that fever has a significant negative effect on recovery time, causing a 21% increase in days needed to recover among patients with fever compared to patients without fever (HR = 0.79, 95% CI: 0.6 to 0.97; $p = 0.03$). Patients with tachypnoea recovered 44% slower than non-

tachypnoea patients (HR = 0.56, 95% CI: 0.38 to 0.82; p = 0.003). Using Kaplan-Meier analysis, we discovered that immunocompromised status did not affect hospitalisation time (3.5 for immunocompromised patients versus 4 for non-immunocompromised patients). However, because the number of observations in the two groups is so different, we can't draw any firm conclusions (HR = 0.93, 95% CI = 0.59 to 1.48, p = 0.69).

4.2.8 Vaccination among hospitalised children

COVID-19 vaccinations became available in Hungary for children aged 12-17 years on June 14, 2021, and for children aged 5-11 years on December 15, 2021. This research spans the months of March 2020 to December 2021. This helps to explain why just 1.55% of the children in the research (5 of 323) received messenger RNA vaccinations. It is crucial to note that 78% of the children in the research were under the age of 12 and would have received the vaccination by mid-December 2021. Because of the few vaccinated children, statistics could not be used to examine how vaccination influenced various clinical markers.

5. DISCUSSION

5.1 Discussion on survey of diagnosis of Epstein-Barr and cytomegalovirus infections using decision trees in paediatric tonsillopharyngitis

Infections of the respiratory tract caused by various viral and bacterial agents are the second most common cause of morbidity in children and adults. Although most upper respiratory tract infections are caused by viruses, a significant proportion of patients are treated with antibiotics. This is due, in part, to doctors' concern about secondary bacterial infection and, in part, to the difficulty of distinguishing viral from bacterial infections based on symptoms. Antimicrobial resistance monitoring research studies in Europe revealed rising resistance rates among Gram-negative and Gram-positive bacteria, with higher levels of resistance found in Europe's south and east. Increasing resistance levels underline the importance of monitoring and avoiding unnecessary antibiotic therapy. Classification algorithms based on machine learning are increasingly employed in clinical research to help making diagnostic decisions.

The classification tree is one such approach that provides an easy way to create flowchart-based categorisation. In this work, we first compared the symptoms and blood test parameters of EBV- and/or CMV-infected patients in the cohort investigated to the non-EBV/CMV group and then utilised the decision tree to develop a model that separates EBV/CMV patients from non-EBV/CMV patients. This research aims to avoid unwarranted antibiotic use in children with tonsillitis. In our study, serological tests showed that 37.60% of patients were infected with EBV

or CMV (30.17% EBV, 2.48% CMV and 4.96% CMV or EBV). These findings are consistent with previous research, which estimated the incidence of EBV in children with tonsillitis ranging from 9% to 58%. Bacteriological culture was also performed on the cohort. Only 2.48% of the overall group tested positive for GAS. This is slightly lower than the stated incidence of 15-30% in the literature, but it is consistent with studies that highlight the overestimation of GAS in tonsillitis.

Following the formation of the EBV/CMV and non-EBV/CMV groups based on the serological test, the different clinical parameters were explored. The findings revealed that lymphadenopathy, hepatosplenomegaly, exudation, and sore throat are the more common symptoms of the EBV/CMV group. Furthermore, we discovered that cough and nasal discharge were much less common in the EBV/CMV group vs the non-EBV/CMV group. This conclusion is consistent with the literature. We confirmed that the mean levels of GOT and GPT were considerably more significant in EBV/CMV patients, while the mean levels of CRP were lower on average (Figure 1). Several investigations have found that CRP levels are higher in bacterial than in viral upper respiratory tract infections. However, our results cannot be fully interpreted in this light, as the non-EBV/CMV group does not only include cases of bacterial infection. As with CRP, the significantly higher mean age in the EBV/CMV group cannot be compared to the published literature.

Decision trees based on symptoms and blood test data were created using univariable statistical methods to investigate the correlations between different variables in the two groups, EBV/CMV and non-EBV/CMV. Not surprisingly, the average classification performance of the blood test-based tree outperformed the symptom-based tree (Figure 2, Figure 3). Our experimental design is complicated to compare with these findings because we did not attempt to discriminate between viral and bacterial infections, but rather one form of viral infection from all others. However, this symptom variance may also explain why it is challenging to effectively differentiate tonsillopharyngitis caused by a specific virus from other pathogens.

 The blood test-based approach, on the other hand, produced better outcomes. The classification effectiveness of the train and test sets did not change, indicating that the differences seen in the train set between the EBV/CMV and non-EBV/CMV groups are generalisable and widely applicable. These variations are mostly related to GPT and CRP levels. According to the tree branches, GPT levels above 34 U/L are specifically characteristic of the EBV/CMV group. The most non-EBV/CMV patients had CRP levels of more than 4.60 mg/L and GPT levels less than 26 U/L. After GPT was found to be such an essential variable in the decision tree, the question was raised whether it could be an applicable marker of EBV/CMV infection. Based on ROC analysis,

GPT might be a well-performing stand-alone marker for the EBV/CMV group, as other research has already shown. However, it is not specific enough: the specificity of the blood test-based tree was 14.48% greater than that of GPT alone (90.3% vs 75.82%). After discovering that the blood test-based decision tree had the most excellent classification efficiency, we investigated how much antibiotic usage may be decreased in the test data set using the tree. To accomplish this, we compared the actual number of EBV/CMV patients receiving antibiotic therapy to the predicted number resulting from the blood test-based tree. The findings demonstrated that unnecessary antibiotic use might be reduced by 66.66%.

In paediatric practice, a rapid antigen detection test (RADT) to prove the presence of GAS is recommended because of its ease of use, short turnaround time, and high specificity. In situations of negative RADT, bacteriological culture is performed to confirm the diagnosis. Most general practitioners lack a diagnostic test to discriminate between viral and bacterial pharyngitis, resulting in antibiotic prescriptions in more than 80% of cases. Clear advice about using culture, serology, and novel diagnostic technologies is essential for reducing needless antibiotic usage. In the lack of a single accurate and rapid differential diagnostic approach, clinical and laboratory results remain crucial in distinguishing between viral and bacterial infections. Although we could not differentiate between viral and bacterial infections in this study, we provide a blood test-based decision tree that may identify EBV/CMV infections with high specificity, decreasing needless antibiotic therapy in children with tonsillopharyngitis.

5.2 Discussion on survey of epidemiology and clinical features of SARS-CoV-2 infection in hospitalised children across four waves from March 2020 to December 2021

This study examined data obtained at the University of Szeged in South Hungary between March 2020 and December 2021. Unlike most original research studies published so far, we had the opportunity to review data from four waves. During the research period, 358 SARS-CoV-2-positive children aged 0 to 18 years were hospitalised at the University of Szeged's Department of Paediatrics. Compared to our time series data, the distribution of the number of cases across time is similar to WHO European coronavirus statistics. Children in the age groups evaluated were more exposed to infection during the fourth wave, caused by the Delta variant (B.1.617.2). Even though a vaccine against the virus was available, just five children received immunisations in the fourth wave

Our findings on the occurrence of the underlying condition are consistent with previous research (Table 3). A higher rate of hospitalised children under the age of one year can be observed,

indicating a greater susceptibility of this age group of patients to symptomatic SARS-CoV-2 infection. However, the higher hospitalisation rate may also be explained by parents seeking medical advice for these children more frequently than older children. Our age distribution is consistent with these results (Figure 5C).

Our investigation's most prevalent COVID-19 symptoms were fever, cough, and nasal discharge. A previously published study found that neurologic symptoms can develop in up to 22% of 1695 hospitalised children and adolescents with severe COVID-19. To these findings, our data indicate a low prevalence of neurological symptoms, with most cases occurring after age 12 (Figure 7). This might be explained by the fact that some subjective symptoms, such as dizziness, loss of smell and taste, and headache, are more difficult to quantify in younger children with less self-expression. In line with a prior study, headache was shown to be our study's most prevalent neurological symptom (Figure 6). In contrast to an international study on neonates, which found that gastrointestinal symptoms and poor oral intake are more common, we found no difference in the distribution of typical symptoms in neonates. Instead, up to 1 month to 1 year of age, a more significant proportion of lower oral intake and diarrhoea is noted (Figure 7). Dehydration was the most prevalent complication in our study, accounting for one of the leading causes of hospitalisation. Dehydration can be caused by virus-induced gastroenteritis, a lack of oral intake due to fever and sore throat, and cough-related vomiting (Figure 6,7). Laryngitis was the fourth most prevalent consequence in our research, involving 8.3% of patients, most toddlers (23.6%) (Figure 6,7). One patient needed intubation and mechanical ventilation due to COVID-induced significant upper airway obstruction. A study of 186 cases from the United States found that a minority of children were symptomatic before the start of MIS-C (Multisystem Inflammatory Syndrome in Children), with a median gap of 25 days between the beginning of COVID-19 symptoms and the onset of MIS-C (60). This explains why we only had two MIS-C patients throughout acute disease (Figure 7). Both patients presented at the second COVID wave. One of them required four days in critical care with cardiac support (Figure 6,7).The most prevalent consequence in newborns was lower respiratory tract inflammation (bronchitis, bronchiolitis), but dehydration was more common in infancy (Figure 6). Only 3% of the patients in our research were admitted to the ICU. Although symptoms were most manifested in the fourth wave with the highest prevalence on average, it did not cause an increase in the proportion of patients requiring ICU (Figure 6).

The observed mortality rate accorded with the findings of other investigations. A case series assessment of publications published up to March 25, 2021. found that the mortality rate among

children with COVID-19 was 0-0.69. We found symptoms and complications, such as fever and tachypnoea, that may have a significant impact on hospitalisation time. However, the negative impact of fever and tachypnoea on hospitalisation time in the present study is smaller. These findings are consistent with prior published studies, which identified fever as one of the most significant negative prognostic factors for recovery time and tachypnoea as a negative predictor of survival and recovery time.

Our clinical research has several limitations. First, there are limitations of retrospective data collection. Comparisons and conclusions are challenging without extensive research spanning numerous waves and a longer COVID-19 timeframe. Furthermore, because we only analysed data from SARS-CoV-2-positive patients who were hospitalised, any further COVID-related problems were excluded from our analysis. Furthermore, SARS-CoV-2 mutations weren't identified in all the patients. Our 22-month retrospective research gives thorough data to help us better understand COVID-19 illness in children. Newer and newer variants of the virus continue to produce varying severity-related diseases. Based on our regional statistics from South-Eastern Hungary, children are characterised by a milder course, lower hospitalisation rates, shorter hospital stays, and reduced death rates, which we expect to remain in future waves.

6. CONCLUSION

6.1 Conclusion of survey of diagnosis of Epstein-Barr and cytomegalovirus infections using decision trees in paediatric tonsillopharyngitis

The first topic of our research is the investigation of paediatric tonsillopharyngitis. We conclude from analysing the results of EBV/CMV serological testing among patients that, while identifying the infection by symptoms is challenging, a model with adequate specificity can be created using blood test parameter

- Considering retrospective evaluation of age, laboratory data, and clinical symptoms, a decision-tree model has been constructed that is simply adaptable in clinical practice to distinguish EBV/CMV-infected patients from infections caused by other pathogens.
- This method can also prevent unnecessary antibiotic usage in upper respiratory tract infections.

6.2 Conclusion of survey of epidemiology and clinical features of SARS-CoV-2 infection in hospitalised children across four waves from March 2020 to December 2021

Our second retrospective study among hospitalised patients with SARS-CoV2 infection provides essential insights to understand the pandemic better.

- More than half (55.6%) of the patients are under the age of three, and one-third (33.8%) are under the age of one year. The fourth wave significantly impacting the Hungarian kid population.

- Appearance of various clinical signs varies with age. Cough, fever, and laryngitis are more common in children under 12. Bronchitis and bronchiolitis have also been identified as frequent problems in children under the age of one month. Virus strains that caused infection early in the COVID era were associated with lower mortality rates and complications in children than adults.

- The third wave featured a higher incidence of gastrointestinal issues, whereas respiratory symptoms dominated the fourth wave.

- We have observed slightly different symptoms and complications in immunocompromised patients. Still, we cannot draw firm conclusions and obtain reliable p-values due to the small number of cases.

- Our data indicate that hospitalisation length is unrelated to age but that specific symptoms (fever and tachypnoea) are linked with prolonged hospitalisation.

This dissertation will fill the gap in knowledge in paediatrics about the epidemiology of COVID-19, as novel viral mutations continue to cause infection nowadays.

The machine learning model for reducing antibiotic use and the knowledge we have gained on COVID-19 infection could help us in the fight against antibiotic resistance in the long term. Considering that, as far as we know, deaths caused by resistant strains of bacteria due to unwarranted antibiotic use are one of the most threatening health problems we face, the results of this research could provide a valuable tool in this seemingly impossible battle.