

Diagnostyka zapalenia wyrostka robaczkowego u dzieci w oparciu o dane kliniczne i laboratoryjne

Vanessa Wiktorja Stabla 

Politechnika Białostocka, Wydział Mechaniczny

e-mail: vanessa.stabla.20.20@wp.pl

Małgorzata Zdrodowska 

Politechnika Białostocka, Wydział Mechaniczny

e-mail: m.zdrodowska@pb.edu.pl

DOI: 10.65731/az-2026-0102

Streszczenie

Ostre zapalenie wyrostka robaczkowego jest najczęstszą chorobą chirurgiczną u dzieci, stanowiąc jedną trzecich hospitalizacji z powodu bólu brzucha w tej grupie wiekowej [Dadeh, Puitong, 2021; Marcinkevičs i in., 2023]. Pomimo dostępnych skal diagnostycznych i postępów w obrazowaniu, rozpoznanie u dzieci wciąż stanowi istotne wyzwanie kliniczne ze względu na nietypowy obraz objawów oraz ograniczoną możliwość komunikacji z młodym pacjentem [Arredondo Montero i in., 2023; Dadeh, Puitong, 2021]. Identyfikacja wiarygodnych predyktorów oraz optymalizacja strategii postępowania, w tym różnicowanie między leczeniem operacyjnym a zachowawczym, może przyczynić się do zmniejszenia odsetka powikłań i poprawy skuteczności terapii [Marcinkevičs i in., 2021; Marcinkevičs i in., 2023]. Celem pracy była analiza zależności pomiędzy występowaniem ostrego zapalenia wyrostka robaczkowego u dzieci a wybranymi zmiennymi klinicznymi, laboratoryjnymi i demograficznymi. Uwzględniono objawy kliniczne, parametry zapalne, wyniki badań ultrasonograficznych oraz wartości wybranych skal diagnostycznych. Analizę przeprowadzono na zbiorze danych „Regensburg Pediatric Appendicitis” dostępnym w repozytorium Zenodo. Dane obejmowały informacje demograficzne (wiek, płeć, BMI), objawy kliniczne (m.in. migracja bólu, tkliwość w prawym dole biodrowym, nudności, wymioty), wyniki badań laboratoryjnych (leukocytoza, CRP, neutrofilia) oraz parametry ultrasonograficzne (średnica wyrostka, obecność wolnego płynu). Uwzględniono także wartości skal diagnostycznych (Alvarado Score, Pediatric Appendicitis Score) oraz zmienne wynikowe, takie jak rozpoznanie i sposób leczenia.

Słowa kluczowe

ostre zapalenie wyrostka robaczkowego, pediatria, diagnostyka, Alvarado Score, Pediatric Appendicitis Score

Wstęp

Ostre zapalenie wyrostka robaczkowego stanowi najczęstszą chorobę chirurgiczną w populacji pediatrycznej, a szczyt zachorowań przypada na przedział wiekowy 9-12 lat. Rozpoznanie choroby u dzieci jest szczególnie problematyczne, ponieważ wielu pacjentów przejawia objawy i oznaki przypominające inne przyczyny bólu brzucha, a młodzi pacjenci często nie wykazują klasycznych cech klinicznych obserwowanych u dorosłych. Opóźniona diagnoza zwiększa ryzyko perforacji wyrostka, co prowadzi do wyższego odsetka powikłań pooperacyjnych, wydłużenia hospitalizacji oraz wzrostu kosztów leczenia [Arredondo Montero i in., 2023; Dadeh, Puitong, 2021; Marcinkevičs i in., 2021; Marcinkevičs i in., 2023].

W odpowiedzi na te trudności opracowano szereg skal punktowych oraz markerów laboratoryjnych, które mają na celu zwiększenie dokładności diagnostycznej i ograniczenie niepotrzebnych interwencji chirurgicznych. Jednym z takich narzędzi jest Pediatric Appendicitis Score, który uwzględnia między innymi: anoreksję, nudności lub wymioty, gorączkę, migrację bólu i wiele innych dolegliwości. Uzyskany wynik pozwala na zmniejszenie ryzyka i podjęcie decyzji o dalszej diagnostyce obrazowej lub leczeniu operacyjnym. Równoległe z oceną kliniczną kluczową rolę odgrywają badania laboratoryjne. Integracja danych klinicznych, laboratoryjnych i ultrasonograficznych z wykorzystaniem metod uczenia maszynowego stwarza możliwość poprawy predykcji sposobu leczenia u dzieci z podejrzeniem ostrego zapalenia wyrostka robaczkowego [Arredondo Montero i in., 2023; Dadeh, Puitong, 2021; Marcinkevičs i in., 2021; Marcinkevičs i in., 2023].

Głównym celem niniejszej analizy była identyfikacja kluczowych czynników klinicznych, laboratoryjnych oraz ultrasonograficznych mających związek z diagnostyką i sposobem leczenia ostrego zapalenia wyrostka robaczkowego u dzieci. Badanie oparto na retrospektywnej analizie danych pacjentów pediatrycznych z podejrzeniem zapalenia wyrostka robaczkowego, obejmującej dane demograficzne, objawy kliniczne, parametry laboratoryjne, wyniki badań ultrasonograficznych oraz wartości skal diagnostycznych Alvarado i Pediatric Appendicitis Score (PAS).

W analizie zastosowano metody statystyczne oraz algorytmy uczenia maszynowego. Zależności między zmiennymi jakościowymi, a wybraną zmienną „Management”, określającą sposób postępowania terapeutycznego, oceniono za pomocą testu

chi kwadrat (χ^2), przyjmując poziom istotności statystycznej $p < 0,05$. Obliczenia statystyczne przeprowadzono w programie Statistica 13.3. Dodatkowo wykorzystano klasyfikatory dostępne w środowisku Weka.

1. Przegląd literatury

Literatura opisuje ostre zapalenie wyrostka robaczkowego jako najczęstszą chorobę chirurgiczną u dzieci, której rozpoznanie stanowi istotne wyzwanie kliniczne ze względu na niespecyficzny obraz objawów oraz trudności w badaniu przedmiotowym w populacji pediatrycznej. W dostępnych publikacjach podkreśla się znaczenie skal punktowych, takich jak Pediatric Appendicitis Score (PAS) i skala Alvarado (skale mają na celu ocenę prawdopodobieństwa występowania ostrego zapalenia wyrostka robaczkowego w grupie wiekowej 3-18 lat). Coraz więcej prac wykorzystuje metody uczenia maszynowego do jednoczesnej predykcji diagnozy, sposobu leczenia (operacyjnego lub zachowawczego) oraz ciężkości choroby, osiągając wysoką skuteczność przewidywania.

Temat diagnostyki zapalenia wyrostka robaczkowego w grupie pediatrycznej jest licznie podejmowany. Statystycznej analizie podjął się między innymi Parveen K., Avabratha Kadke S. oraz Shetty K. w swoim artykule zatytułowanym „Pediatric appendicitis score in the diagnosis of childhood appendicitis: a validation study”. Przeprowadzono w nim prospektywne badanie walidacyjne skali PAS u 60 dzieci w wieku od 4 do 16 lat. Analiza statystyczna obejmowała porównanie częstości występowania poszczególnych objawów i parametrów laboratoryjnych pomiędzy grupami z różnymi wynikami PAS z wykorzystaniem testu chi kwadrat (χ^2), gdzie za istotne statystycznie uznano wartości p wynoszące odpowiednio 0,001 dla anoreksji, 0,004 dla gorączki, 0,023 dla migracji bólu, 0,000 dla bolesności przy kaszlu, 0,000 dla leukocytozy oraz 0,007 dla neutrofilii. Obliczono również czułość (57,6%), swoistość (85,2%), dodatnią wartość predykcyjną (75%), ujemną wartość predykcyjną (72,5%) oraz dodatni i ujemny wskaźnik wiarygodności (odpowiednio 3,92 i 0,5) dla skali PAS. Dane analizowane w tym badaniu obejmowały wiek, płeć, obecność anoreksji, nudności lub wymiotów, gorączki, migracji bólu, tkliwości w prawym dolnym kwadrancie, bolesności przy kaszlu, leukocytozy, neutrofilii, wyniki ultrasonografii (obecność zapalenia wyrostka) oraz wynik badania histopatologicznego.

Pierwszym objawem zapalenia wyrostka robaczkowego jest ból brzucha dziecka. Dadeh A. A. oraz Puitong K. w swoim artykule z 2021 roku przeprowadzili badanie na 1043 dzieciach z tym właśnie objawem. Autorzy zidentyfikowali za po-

mocą wieloczynnikowej regresji logistycznej cztery niezależne czynniki predykcyjne rozpoznania zapalenia wyrostka: stopniowe narastanie bólu brzucha, tkliwość w prawym dolnym kwadrancie, obecność cech podrażnienia otrzewnej oraz odsetek neutrofilów powyżej 75%, a opracowany nomogram osiągnął pole pod krzywą ROC wynoszące 0,963. Dane analizowane w tym badaniu to wiek, płeć, masa ciała, wzrost, wskaźnik masy ciała, czas trwania bólu, ból okołopępkowy, nasilenie bólu przy ruchu, migracja bólu do prawego dolnego kwadrantu, stopniowe narastanie bólu, nudności i wymioty, gorączka, tachykardia, tkliwość w prawym dolnym kwadrancie, cechy podrażnienia otrzewnej, liczba białych krwinek, odsetek neutrofilów, stosunek neutrofilów do limfocytów, średnia objętość płytek krwi, obecność krwinek białych lub czerwonych w moczu, wyniki ultrasonografii i tomografii komputerowej oraz diagnoza końcowa.

W badaniu Arredondo Montero J., Pascual C., oraz współautorów (2023) udało się opracować indeks BIDIAP uwzględniający trzy niezależne predyktory: kaliber wyrostka robaczkowego $\geq 6,9$ mm (4 punkty), wskaźnik ogólnoustrojowej odpowiedzi zapalnej SII ≥ 890 (3 punkty) oraz obecność podrażnienia otrzewnej (2 punkty), przy czym indeks ten osiągnął pole pod krzywą ROC 0,97, a przy wartości odcięcia 4 punktów uzyskał czułość 98,98% i swoistość 77,78%. Dane analizowane w tym badaniu to wiek, płeć, liczba godzin ewolucji bólu, gorączka w domu, liczba stolców biegunkowych, objawy ze strony układu moczowego, liczba epizodów wymiotnych, hiporeksja, liczba leukocytów, neutrofilów i płytek krwi, wskaźnik SII (obliczony jako iloczyn liczby neutrofilów i płytek podzielony przez liczbę limfocytów), kaliber wyrostka robaczkowego w badaniu USG, obecność złoju kałowego w świetle wyrostka, obecność zapalenia węzłów chłonnych krezkowych, obecność podrażnienia otrzewnej oraz stężenie interleukiny 6, białka C reaktywnego i pentraksyny 3.

W literaturze [Marcinkevics R. i Reis Wolfertstetter P., 2021] wykazano również, że zastosowanie algorytmów uczenia maszynowego pozwoliło na wykazanie wysokiej skuteczności predykcji diagnozy (pole pod krzywą precyzji czułości 0,94), sposobu leczenia (0,92) oraz ciężkości choroby (0,70), przy czym za najważniejsze zmienne uznano średnicę wyrostka, widoczność wyrostka w USG, podrażnienie otrzewnej, liczbę białych krwinek, odsetek neutrofilów oraz stężenie białka C-reaktywnego. Dane analizowane w tym badaniu to 38 zmiennych predykcyjnych, w tym wiek, płeć, wzrost, masa ciała, wskaźnik masy ciała, wynik Alvarado, wynik PAS, obecność podrażnienia otrzewnej, migracja bólu, tkliwość w prawym dolnym kwadrancie, tkliwość odbita, bolesność przy kaszlu, objaw psoas, nudności i wymioty, anoreksja, temperatura ciała, dysuria, nieprawidłowe stolce, liczba białych krwinek, odsetek neutrofilów, stężenie białka C reaktywnego, ketony w moczu, krwinki czerwone i białe w moczu, widoczność wyrostka w USG, kaliber wyrostka, wolny płyn

w jamie otrzewnej, nieregularne warstwy wyrostka, objaw tarczy, perfuzja wyrostka, reakcja tkanki otaczającej, patologiczne węzły chłonne, zapalenie węzłów chłonnych krezkowych, pogrubienie ściany jelita, ileus, koprostaza, meteoryzm oraz zapalenie jelit.

2. Metodyka badań

Badanie miało charakter analizy statystycznej oraz eksploracyjnej analizy danych z wykorzystaniem metod uczenia maszynowego, której celem była ocena zależności pomiędzy wybranymi zmiennymi klinicznymi, laboratoryjnymi i ultrasonograficznymi, a sposobem postępowania terapeutycznego u dzieci z podejrzeniem ostrego zapalenia wyrostka robaczkowego, przy czym w ramach analizy zastosowano klasyfikatory dostępne w środowisku Weka w celu oceny skuteczności predykcji zmiennej „Management” określającej sposób leczenia pacjenta. Wykorzystano klasyfikatory, takie jak [Zdrodowska i Dardzińska-Głębocka, 2022; Turiac, i Zdrodowska, 2022; Zdrodowska, 2019; Zdrodowska i Kasperczuk, 2023]:

- RBFClassifier, który opiera się na sieci neuronowej z radialnymi funkcjami bazowymi (Radial Basis Function). W modelu tym każda jednostka ukryta realizuje funkcję Gaussa, której aktywacja maleje wraz z odległością od centrum. Sieć RBF składa się z trzech warstw: wejściowej, ukrytej z nieliniowymi funkcjami radialnymi oraz wyjściowej liniowej, co pozwala na efektywne aproksymowanie dowolnych funkcji ciągłych przy stosunkowo krótkim czasie uczenia.
- LibSVM, który jest implementacją maszyny wektorów nośnych (Support Vector Machine) z wykorzystaniem jądra radialnego lub innych funkcji jądrowych. Algorytm ten konstruuje optymalną hiperpłaszczyznę separującą klasy poprzez maksymalizację marginesu. LibSVM rozwiązuje zadanie programowania kwadratowego z wykorzystaniem metody sekwencyjnej minimalizacji optymalizacji (SMO), co zapewnia wysoką skuteczność w problemach nieliniowych.
- kNN (k najbliższych sąsiadów), który klasyfikuje nowy przypadek na podstawie odległości do k najbliższych przykładów uczących. Odległość obliczana jest w różny sposób w zależności od wybranej metryki.
- J48 jest implementacją algorytmu generującego drzewo decyzyjne. W każdym węźle algorytm wybiera atrybut o najwyższym znormalizowanym przyroście informacji i dzieli zbiór na podzbiory. Proces zatrzymuje się, gdy podzbiór zawiera przypadki należące do jednej klasy lub gdy dalszy podział nie poprawia jakości klasyfikacji.

- REPTree (Reduced Error Pruning Tree), który buduje szybkie drzewo decyzyjne z wykorzystaniem przycinania przez redukcję błędu. Algorytm stosuje metodę podziału opartą na wariancji lub entropii, a następnie przycina gałęzie, które nie zmniejszają błędu na zbiorze walidacyjnym, co pozwala uzyskać prostsze drzewo bez znaczącej utraty dokładności.
- SimpleCart (CART – Classification and Regression Trees), który tworzy binarne drzewo decyzyjne, w którym każdy węzeł dzielony jest tylko na dwa podobozy. Algorytm wykorzystuje miarę nieczystości Gini do wyboru najlepszego podziału. Drzewo jest następnie przycinane za pomocą kosztowo złożeniowego przycinania (cost complexity pruning), co pozwala na uzyskanie modeli łatwych do interpretacji klinicznej.
- PART (Partial Decision Trees), który łączy strategię dziel i zwyciężaj z budową częściowych drzew decyzyjnych. Tworzy reguły poprzez konstruowanie drzewa dla aktualnego zbioru, następnie wybiera najlepszą ścieżkę jako regułę, usuwa pokryte nią przypadki i powtarza proces.
- RandomTree, który buduje pojedyncze drzewo decyzyjne na podstawie losowo wybranego podzbioru atrybutów w każdym węźle. W przeciwieństwie do lasów losowych, które agregują wiele drzew, RandomTree generuje jedno drzewo bez procesu głosowania. Algorytm nie przeprowadza przycinania, co może prowadzić do przeuczenia, ale przy odpowiedniej liczbie atrybutów w próbce pozwala na szybką eksplorację danych.
- NaiveBayes, który opiera się na twierdzeniu Bayesa z założeniem niezależności warunkowej atrybutów względem klasy. Mimo że założenie to jest często nieprawdziwe w rzeczywistych danych, algorytm osiąga zaskakująco dobrą skuteczność w wielu zadaniach klasyfikacji medycznej. Oblicza prawdopodobieństwo a posteriori dla każdej klasy i wybiera klasę o największej wartości.
- MultilayerPerceptron, który jest sztuczną siecią neuronową o architekturze feed forward składającą się z warstwy wejściowej, jednej lub więcej warstw ukrytych z nieliniowymi funkcjami aktywacji oraz warstwy wyjściowej. Uczenie odbywa się za pomocą algorytmu wstecznej propagacji błędów (backpropagation), który minimalizuje błąd średniokwadratowy poprzez iteracyjną modyfikację wag. Sieć ta potrafi modelować bardzo złożone zależności, jednak wymaga znacznych zasobów obliczeniowych i dużej liczby przykładów uczących.

Następnie przeprowadzono analizę zależności pomiędzy tą zmienną, a pozostałymi zmiennymi z wykorzystaniem testu chi kwadrat (χ^2) w programie Statistica 13.3 (badanie przeprowadzono na poziomie istotności $p < 0,05$ (przy czym wyniki

o niższej wartości p uznawano za istotne statystycznie)), a dla zmiennych wykazujących istotne statystycznie powiązania ponownie przeprowadzono analizy klasyfikacyjne z użyciem wybranych algorytmów uczenia maszynowego, co umożliwiło ocenę ich wartości diagnostycznej oraz przydatności w przewidywaniu sposobu leczenia pacjentów. Analizę przeprowadzono na podstawie bazy danych obejmującej 423 pacjentów młodocianych, u których stwierdzono zapalenie wyrostka robaczkowego [Marcinkevičs i in, 2023]. Każdemu pacjentowi przypisano 58 zmiennych, w skład których wchodziły:

- Dane pacjenta: wiek, wzrost, masa ciała, BMI;
- Długość hospitalizacji - liczba dni pobytu w szpitalu;
- Skala Alvarado (AS) - wynik punktowy w skali Alvarado służącej do oceny prawdopodobieństwa zapalenia wyrostka robaczkowego;
- Pediatriańska skala zapalenia wyrostka (PAS) - wynik punktowy w pediatriańskiej skali diagnostycznej zapalenia wyrostka robaczkowego;
- Otrzewnowe napięcie mięśniowe / obrona mięśniowa: brak, miejscowa, uogólniona (obecność napięcia mięśni brzucha świadczącego o stanie zapalnym otrzewnej);
- Migracja bólu: nie/tak (przemieszczenie bólu z nadbrzusza do prawego dołu biodrowego);
- Tkliwość w prawym dolnym kwadrancie brzucha (RLQ): nie/tak (bolesność palpacyjna prawego dołu biodrowego);
- Objaw Blumberga po stronie przeciwnej: nie/tak (ból po zwolnieniu ucisku po przeciwnej stronie brzucha);
- Objaw Blumberga po stronie tej samej: nie/tak (ból po zwolnieniu ucisku po tej samej stronie brzucha);
- Ból przy kaszlu: nie/tak
- Objaw mięśnia lędźwiowego (objaw psoas): ujemny/dodatni (ból wywołany prostowaniem stawu biodrowego);
- Nudności / wymioty: nie/tak;
- Utrata apetytu: nie/tak;
- Temperatura ciała [$^{\circ}\text{C}$];
- Dysuria: nie/tak (ból lub trudności podczas oddawania moczu);
- Charakter stolca: prawidłowy, biegunka, zaparcie;
- Liczba leukocytów (WBC) - liczba białych krwinek we krwi;
- Liczba erytrocytów (RBC) - liczba czerwonych krwinek we krwi;
- Hemoglobina - stężenie hemoglobiny we krwi;
- RDW - wskaźnik zróżnicowania objętości erytrocytów;
- Liczba płytek krwi;

- Neutrofile: procentowy udział neutrofili we krwi;
- Neutrofilia: nie/tak (zwiększony odsetek neutrofili ($\geq 75\%$));
- Białko C-reaktywne (CRP) - marker stanu zapalnego;
- Ciała ketonowe w moczu: brak, +, ++, +++ (obecność ketonów w moczu);
- Erytrocyty w moczu: negatywny, +, ++, +++ (obecność krwi w moczu);
- Leukocyty w moczu: brak, +, ++, +++ (obecność leukocytów w moczu);
- Wykonano USG: nie/tak (informacja o wykonaniu badania ultrasonograficznego jamy brzusznej);
- Widoczność wyrostka robaczkowego w USG: nie/tak;
- Średnica wyrostka robaczkowego: wartość ciągła – maksymalna średnica wyrostka w mm;
- Wolny płyn w jamie otrzewnej: nie/tak;
- Reakcja tkanek otaczających: nie/tak (cechy zapalenia w tkankach wokół wyrostka);
- Rozpoznanie wstępne: tekst – podejrzenie diagnostyczne postawione przy przyjęciu;
- Rozpoznanie końcowe: brak zapalenia wyrostka, zapalenie wyrostka – ostateczne rozpoznanie pacjenta;
- Sposób leczenia: zachowawcze, pierwotnie operacyjne, wtórnie operacyjne – sposób postępowania terapeutycznego;
- Stopień ciężkości: brak / niepowikłane, powikłane – klasyfikacja ciężkości zapalenia wyrostka robaczkowego.

3. Wyniki badań

Pierwszym etapem analizy była ocena skuteczności predykcji za pomocą dostępnych w oprogramowaniu Weka klasyfikatorów. Analizowane były wyniki między innymi, takie jak: dokładność [%] (ogólna skuteczność klasyfikatora), TPRate [-] (oznaczającego poprawnie wykrytych przypadków danej klasy), precyzja [-] (która określa, jaki odsetek przewidzianych przez model przypadków był rzeczywiście poprawny), pole pod krzywą ROC [-] (ocena zdolności modelu do rozróżniania klas) czas generacji modelu [s] oraz macierz pomyłek (pokazująca liczbę poprawnych i błędnych klasyfikacji dla każdej klasy). Zastosowano podstawowe wartości klasyfikatorów, a wynik analizy przedstawiono w tabeli 1.

Tab. 1. Wyniki analizy klasyfikatorów zmiennej Management

zastosowany klasyfikator	dokładność [%]	TPRate [-]	pole pod krzywą ROC [-]	czas generacji [s]
RBFClassifier	91,7258	0,917	0,947	0,15

Tab. 1. Wyniki analizy klasyfikatorów zmiennej Management

zastosowany klasyfikator	dokładność [%]	TPRate [-]	pole pod krzywą ROC [-]	czas generacji [s]
LibSVM	83,6879	0,837	0,500	0,07
kNN / Eucliden	89,3617	0,894	0,790	0
kNN/Manhattan	90,3073	0,903	0,794	0
J48	98,1087	0,981	0,962	0,02
REPTree	96,4539	0,965	0,943	0,01
SimpleCart	95,9811	0,96	0,920	0,12
PART	98,3452	0,983	0,987	0,01
RandomTree	91,7258	0,917	0,893	0
NaiveBayers	91,2530	0,913	0,945	0
MultiplayerPercepction	92,1986	0,922	0,933	4,92

Źródło: opracowanie własne na podstawie [Marcinkevičs i in., 2023].

Uzyskane wyniki wskazują na wyraźne różnice skuteczności pomiędzy zastosowanymi klasyfikatorami w predykcji zmiennej „Management”. Najwyższą skuteczność klasyfikacji uzyskał algorytm PART, osiągając dokładność na poziomie 98,35%, wartość TP Rate równą 0,983 oraz najwyższe pole pod krzywą ROC wynoszące 0,987. Otrzymane wyniki świadczą o bardzo wysokiej zdolności modelu do poprawnego rozróżniania klas oraz niewielkiej liczbie błędnych klasyfikacji. Bardzo zbliżone rezultaty uzyskał również klasyfikator J48, którego dokładność wyniosła 98,11%, przy TP Rate i precyzji równych 0,981 oraz ROC = 0,962. Wskazuje to, że modele oparte na drzewach decyzyjnych charakteryzowały się najwyższą efektywnością w analizowanym zbiorze danych.

Również dobre wyniki osiągnęły klasyfikatory REPTree oraz SimpleCart, których dokładność przekraczała 95%. Modele te uzyskały również wysokie wartości pola pod krzywą ROC (>0,92), co potwierdza ich dobrą zdolność klasyfikacyjną. Nieco niższą, ale nadal wysoką skuteczność wykazały modele RandomTree, Naive-Bayes oraz Multilayer Perceptron, osiągając dokładność na poziomie około 91–92%.

Najslabsze wyniki uzyskał klasyfikator LibSVM, którego dokładność wyniosła 83,69%, a pole pod krzywą ROC osiągnęło wartość 0,500. Taki wynik ROC wskazuje na bardzo ograniczoną zdolność modelu do rozróżniania klas i sugeruje skuteczność zbliżoną do klasyfikacji losowej. Może to świadczyć o niedopasowaniu parametrów modelu SVM do charakterystyki analizowanego zbioru danych lub o niewielkiej separowalności klas przy zastosowanych ustawieniach algorytmu.

Czas generacji modeli, w przypadku większość klasyfikatorów był on bardzo krótki, nieprzekraczający 0,15 s. Wyjątek stanowił Multilayer Perceptron, dla którego czas generacji wyniósł 4,92 s, co było zdecydowanie najwyższą wartością spośród analizowanych modeli.

Ogół wyników wskazuje, że najwyższą skuteczność diagnostyczną w analizie zmiennej „Management” wykazały klasyfikatory oparte na drzewach decyzyjnych, szczególnie PART oraz J48, które łączyły bardzo wysoką dokładność klasyfikacji z wysokimi wartościami ROC i krótkim czasem generacji modelu.

W dalszej części badania przystąpiono do analizy zależności dostępnych zmiennych na wcześniej analizowaną zmienną management. Rozpoczęto od weryfikacji normalności rozkładu. Z racji braku rozkładu normalnego zmiennej „hemoglobin” zdecydowano się na wykorzystania testu U Manna-Whitneya. W przypadku pozostałych zmiennych zastosowano test chi kwadrat (χ^2).

Tab. 2. Wyniki testu chi-kwadrat Pearsona zmiennej Management

zmienna	p	zmienna	p
Age	1	Neutrophil_Percentage	0,357
BMI	1	Neutrophilia	< 0,001
Sex	0,391	RBC_Count	0,513
Height	0,241	Hemoglobin	0,689
Weight	0,185	RDW	0,122
Length_of_Stay	< 0,001	Thrombocyte_Count	0,754
Severity	< 0,001	Ketones_in_Urine	< 0,001
Diagnosis_Presumptive	< 0,001	RBC_in_Urine	0,028
Diagnosis	< 0,001	RBC_in_Urine	0,545
Alvarado_Score	< 0,001	WBC_in_Urine	1
Paediatric_Appendicitis_Score	< 0,001	CRP	< 0,001
Appendix_on_US	0,023	Dysuria	0,756
Migratory_Pain	0,178	Stool	0,482
Lower_Right_Abd_Pain	0,633	Peritonitis	< 0,001
Contralateral_Rebound_Tenderness	< 0,001	Psoas_Sign	0,751
Coughing_Pain	0,365	Ipsilateral_Rebound_Tenderness	< 0,001
Nausea	< 0,001	US_Performed	1
Loss_of_Appetite	0,129	US_Number	0,477

zmienna	p	zmienna	p
Body_Temperature	0,319	Free_Fluids	< 0,001
WBC_Count	0,009		

Źródło: opracowanie własne na podstawie [Marcinkevičs i in., 2023].

W przeprowadzonej analizie wykazano istotną statystycznie zależność pomiędzy długością hospitalizacji (Length_of_Stay) a analizowaną zmienną wynikową ($p < 0,001$). Zaobserwowano, że pacjenci wymagający bardziej zaawansowanego postępowania terapeutycznego lub z cięższym przebiegiem choroby charakteryzowali się dłuższym czasem pobytu w szpitalu.

Istotną zależność stwierdzono również dla sposobu leczenia (Management) oraz stopnia ciężkości choroby (Severity) ($p < 0,001$). Wyniki wskazują, że cięższy przebieg zapalenia wyrostka robaczkowego wiązał się częściej z koniecznością leczenia operacyjnego, szczególnie w przypadkach powikłanych.

Analiza wykazała także silny związek pomiędzy rozpoznaniem wstępnym (Diagnosis_Presumptive) oraz rozpoznaniem końcowym (Diagnosis) a analizowaną zmienną zależną ($p < 0,001$). Sugeruje to wysoką wartość diagnostyczną oceny klinicznej przeprowadzanej podczas przyjęcia pacjenta do szpitala.

W przypadku skal diagnostycznych tj. skali Alvarado (Alvarado_Score) oraz Pediatric Appendicitis Score (Paediatric_Appendicitis_Score) również wykazano istotne zależności statystyczne ($p < 0,001$). Wyższe wartości punktowe tych skal korelowały z większym prawdopodobieństwem wystąpienia ostrego zapalenia wyrostka robaczkowego oraz koniecznością wdrożenia leczenia chirurgicznego. Wyniki te potwierdzają przydatność obu skal jako narzędzi wspomagających proces diagnostyczny.

W analizie parametrów ultrasonograficznych wykazano istotny związek pomiędzy widocznością wyrostka robaczkowego w badaniu USG (Appendix_on_US) a badaną zmienną ($p = 0,023$). Obecność uwidocznionego wyrostka częściej występowała u pacjentów z potwierdzonym stanem zapalnym.

Istotne zależności zaobserwowano również dla objawów klinicznych świadczących o podrażnieniu otrzewnej, takich jak objaw Blumberga po stronie przeciwnej (Contralateral_Rebound_Tenderness; $p = 0,0002$), objaw Blumberga po stronie tej samej (Ipsilateral_Rebound_Tenderness; $p < 0,001$) oraz obecność objawów otrzewnowych (Peritonitis; $p < 0,001$). Wyniki te wskazują, że nasilone objawy otrzewnowe mogą być związane z bardziej zaawansowanym przebiegiem zapalenia wyrostka robaczkowego.

Wykazano także istotną zależność pomiędzy występowaniem nudności i wymiotów (Nausea) a badaną zmienną ($p < 0,001$). Objawy te częściej obserwowano u pacjentów z rozpoznanym ostrym zapaleniem wyrostka robaczkowego.

W zakresie parametrów laboratoryjnych istotne statystycznie zależności dotyczyły liczby leukocytów (WBC_Count; $p = 0,009$), neutrofilii (Neutrophilia; $p < 0,001$) oraz stężenia białka C-reaktywnego CRP ($p < 0,001$). Podwyższone wartości markerów stanu zapalnego były częściej obserwowane u pacjentów z ostrym stanem zapalnym oraz cięższym przebiegiem choroby, co potwierdza ich znaczenie diagnostyczne w ocenie zapalenia wyrostka robaczkowego.

Dodatkowo stwierdzono istotną zależność pomiędzy obecnością erytrocytów w moczu (RBC_in_Urine) a analizowaną zmienną ($p = 0,028$). Wynik ten może wskazywać na możliwość współwystępowania objawów ze strony układu moczowego lub trudności diagnostycznych wynikających z lokalizacji procesu zapalnego.

W badaniach obrazowych wykazano także istotny związek pomiędzy obecnością wolnego płynu w jamie otrzewnej (Free_Fluids), a analizowaną zmienną ($p < 0,001$). Obecność wolnego płynu częściej obserwowano u pacjentów z bardziej nasilonym procesem zapalnym lub powikłanym przebiegiem choroby.

Uzyskane wyniki podkreślają znaczenie kompleksowej oceny klinicznej, laboratoryjnej oraz ultrasonograficznej w diagnostyce ostrego zapalenia wyrostka robaczkowego u dzieci. Szczególnie wysoką wartość diagnostyczną wykazały objawy otrzewnowe, markery stanu zapalnego oraz wyniki skal diagnostycznych, które mogą wspierać podejmowanie decyzji terapeutycznych i poprawiać skuteczność wczesnego rozpoznania choroby.

Następnie przystąpiono do analizy skuteczności predykcji wcześniej wykorzystanych klasyfikatorów dla zmiennych, u których zaobserwowano istotną zależność statystyczną po analizie zmiennej „Management”. Ze względu na rodzaj zmiennych i braku możliwości analizy zmiennych numerycznych za pomocą klasyfikatorów w oprogramowaniu Weka wykonano analizę z pominięciem zmiennych: Length_of_Stay, Alvarado_Score, Paediatric_Appendicitis_Score, WBC_Count oraz CRP. Wyniki analizy przedstawiono w tabeli 3 oraz 4, natomiast spis oznaczeń klasyfikatorów przedstawiono w tabeli 5.

Tab. 3. Wyniki badania zmiennych istotnych statystycznie cz.1

zmianna	C1	C2	C3	C4	C5	C6
Severity	95,508	95,272	94,799	95,745	94,799	94,563
Diagnosis_Presump- tive	89,125	85,579	84,397	85,816	94,09	93,617
Diagnosis	80,615	53,428	71,868	73,995	84,87	88,889
Appendix_on_US	77,541	63,83	70,686	73,05	79,196	75,414
Contralateral_Rebo- und_Tenderness	68,794	64,066	64,303	62,884	70,213	69,976
Nausea	72,813	55,556	66,667	70,449	76,832	75,65
Neutrophilia	90,307	60,993	73,286	79,905	98,582	98,818
RBC_in_Urine	77,327	78,52	71,122	70,644	74,463	77,566
Peritonitis	85,343	81,088	77,541	77,541	83,924	86,052
Ipsilateral_Rebo- und_Tenderness	95,508	95,745	93,853	93,853	93,853	95,745
Free_Fluids	61,702	61,702	57,683	58,156	56,265	62,648

Źródło: opracowanie własne.

Tab. 4. Wyniki badania zmiennych istotnych statystycznie cz.2

zmianna	C7	C8	C9	C10	C11
Severity	96,454	95,508	93,853	94,563	95,508
Diagnosis_Presump- tive	94,563	94,326	87,47	90,307	86,525
Diagnosis	91,253	87,943	78,014	72,104	78,487
Appendix_on_US	80,142	73,995	65,248	67,612	72,813
Contralateral_Rebo- und_Tenderness	70,213	65,485	60,52	58,392	71,158
Nausea	75,414	72,104	75,65	62,411	70,922
Neutrophilia	99,527	98,345	87,707	86,761	93,853
Peritonitis	87,234	78,723	73,995	72,34	81,324
Ipsilateral_Rebo- und_Tenderness	95,745	93,617	91,726	86,052	93,617
Free_Fluids	61,466	56,028	50,591	60,757	53,192

Źródło: opracowanie własne.

Tab. 5. Oznaczenia analizowanych klasyfikatorów

RBFClassifier	C1
LibSVM	C2
kNN / Eucliden	C3
kNN/Manhattan	C4
J48	C5
REPTree	C6
SimpleCart	C7
PART	C8
RandomTree	C9
NaiveBayers	C10
MultiplayerPerception	C11

Źródło: opracowanie własne.

Analiza dokładności klasyfikatorów dla poszczególnych zmiennych wykazała istotne różnice w dokładności pomiędzy konkretnymi badanymi zmiennymi. Najwyższą dokładność klasyfikacji uzyskano dla zmiennej „Neutrophilia”. W przypadku tej zmiennej szczególnie wysoką skuteczność osiągnęły klasyfikatory oparte na drzewach decyzyjnych. Najlepszy wynik uzyskał model SimpleCart (99,53%), następnie REPTree (98,82%), J48 (98,58%) oraz PART (98,35%). Wyniki te wskazują, że zmienna ta stanowi bardzo silny czynnik różnicujący analizowane klasy i może mieć istotną wartość diagnostyczną.

Bardzo wysoką dokładność uzyskano również dla zmiennej „Severity”. Wszystkie klasyfikatory osiągnęły wyniki przekraczające 93%, co wskazuje na dużą stabilność i wysoką przewidywalność tej cechy. Najlepszy rezultat osiągnął model SimpleCart (96,45%), natomiast najniższą skuteczność odnotowano dla RandomTree (93,85%). Niewielkie różnice pomiędzy modelami sugerują, że stopień ciężkości choroby jest zmienną dobrze separującą klasy niezależnie od zastosowanego algorytmu.

W przypadku zmiennej „Ipsilateral_Rebound_Tenderness” również zaobserwowano bardzo wysoką skuteczność klasyfikacji. Najlepsze wyniki osiągnęły klasyfikatory LibSVM, REPTree oraz SimpleCart (95,74%). Dla zmiennej „Diagnosis_Presumptive” najwyższą dokładność uzyskały klasyfikatory SimpleCart (94,56%), PART (94,33%) oraz J48 (94,09%). Modele drzewiaste ponownie wykazały przewagę nad pozostałymi algorytmami. Zmienne „Diagnosis” oraz „Peritoni-

tis” osiągnęły umiarkowanie wysokie wartości dokładności. W przypadku rozpoznania końcowego najwyższą skuteczność uzyskał model SimpleCart (91,25%), natomiast zdecydowanie najslabszy wynik odnotowano dla LibSVM (53,43%). Tak duża różnica pomiędzy klasyfikatorami może wskazywać na złożoną strukturę danych oraz większą skuteczność modeli drzewiastych w wykrywaniu zależności pomiędzy analizowanymi cechami.

Dla zmiennej „Appendix_on_US” najlepszy wynik uzyskał model SimpleCart (80,14%), natomiast najniższą skuteczność ponownie wykazał LibSVM (63,83%). Wyniki sugerują umiarkowaną wartość diagnostyczną widoczności wyrostka robaczkowego w badaniu ultrasonograficznym. W przypadku objawu „Nausea” najwyższą dokładność osiągnął klasyfikator J48 (76,83%), natomiast najniższy wynik uzyskał LibSVM (55,56%). Wskazuje to, że nudności i wymioty mają ograniczoną wartość jako pojedynczy parametr klasyfikacyjny. Dla zmiennej „Contralateral_Rebound_Tenderness” uzyskano relatywnie niższe wartości dokładności. Najlepszy rezultat osiągnął Multilayer Perceptron (71,16%), a najniższy NaïveBayes (58,39%). Wyniki te mogą sugerować większą zmienność kliniczną tego objawu oraz mniejszą zdolność do jednoznacznego różnicowania pacjentów. Zmienna „Free_Fluids” charakteryzowała się najniższą skutecznością klasyfikacji spośród analizowanych parametrów. Pierwszej z nich najwyższą dokładność uzyskał LibSVM (78,52%), podczas gdy większość pozostałych modeli osiągała wyniki w zakresie 64–77%. W przypadku drugiej zmiennej najwyższy wynik uzyskał REPTree (62,65%), natomiast RandomTree osiągnął jedynie 50,59%, co wskazuje na bardzo ograniczoną zdolność klasyfikacyjną tej zmiennej.

4. Dyskusja wyników

Przeprowadzona analiza potwierdziła, że diagnostyka ostrego zapalenia wyrostka robaczkowego u dzieci wymaga jednoczesnego uwzględnienia danych klinicznych, laboratoryjnych i obrazowych, a połączenie klasycznej oceny z metodami uczenia maszynowego może znacząco zwiększyć skuteczność identyfikacji pacjentów wymagających leczenia operacyjnego. W badaniu wykazano istotne statystycznie zależności pomiędzy neutrofilią, podwyższonym stężeniem CRP, obecnością objawów otrzewnowych oraz wynikami skal Alvarado i Pediatric Appendicitis Score, a sposobem leczenia, co pozostaje zgodne z doniesieniami literaturowymi.

Parveen i współpracownicy w prospektywnym badaniu walidacyjnym skali PAS potwierdzili istotność anoreksji, gorączki, migracji bólu, bolesności przy kaszlu, leukocytozy i neutrofilii, osiągając czułość 57,6% i swoistość 85,2%.

Dadeh i Puitong zidentyfikowali stopniowe narastanie bólu, tklliwość w prawym dolnym kwadrancie, cechy podrażnienia otrzewnej oraz neutrofilię powyżej 75% jako niezależne predyktory, a opracowany nomogram osiągnął pole pod krzywą ROC wynoszące 0,963.

Arredondo Montero i zespół opracowali indeks BIDIAP, uwzględniający kaliber wyrostka $\geq 6,9$ mm, wskaźnik SII ≥ 890 oraz podrażnienie otrzewnej, który przy wartości odcięcia 4 punktów uzyskał czułość 98,98% i swoistość 77,78% oraz pole pod krzywą ROC 0,97.

Marcinkevics i Reis Wolfertstetter zastosowali algorytmy uczenia maszynowego u 430 dzieci, wykazując wysoką skuteczność lasów losowych i modeli ze wzmocnieniem gradientowym w predykcji diagnozy (pole pod krzywą precyzji czułości 0,94), sposobu leczenia (0,92) i ciężkości choroby (0,70), przy czym za kluczowe zmienne uznano średnicę wyrostka, widoczność w USG, podrażnienie otrzewnej, liczbę leukocytów, odsetek neutrofilii oraz stężenie CRP.

W niniejszej analizie najwyższą skuteczność klasyfikacji uzyskano dla neutrofilii, sięgającą ponad 99% dla modelu SimpleCart, natomiast klasyfikatory drzewiaste, takie jak PART, J48, REPTree i SimpleCart, wyraźnie przewyższały algorytm LibSVM oraz sieć neuronową Multilayer Perceptron. Zmienne ultrasonograficzne, w tym RBC_in_Urine i Appendix_on_US wykazywały istotne statystycznie zależności, lecz skuteczność klasyfikacji drugiej zmiennej była umiarkowana, co może wynikać z operatorozależności badania USG oraz niespecyficzności wolnego płynu, który występuje także w innych schorzeniach jamy brzusznej. Uzyskane wyniki potwierdzają, że najbardziej wartościowymi czynnikami diagnostycznymi pozostają objawy podrażnienia otrzewnej, neutrofilia, podwyższone stężenie CRP oraz wyniki skal Alvarado i PAS, a zastosowanie algorytmów uczenia maszynowego, zwłaszcza modeli drzewiastych, może stanowić cenne wspomaganie procesu decyzyjnego, jednak nie zastępuje oceny klinicznej lekarza.

Podsumowanie

Wykonanie analizy statystycznej wraz z przeglądem literatury pozwala na zaobserwowanie praktycznego oraz naukowego wpływu badań na rozwój diagnostyki zapalenia wyrostka robaczkowego u dzieci. W praktyce, najważniejsza jest skuteczność diagnostyka, która wymaga integracji trzech grup parametrów: objawów klinicznych, markerów laboratoryjnych oraz wyników ultrasonografii. Żaden pojedynczy parametr nie jest wystarczająco czuły ani swoisty, aby samodzielnie przesądzać o rozpoznaniu lub decyzji terapeutycznej. Skale punktowe, takie jak Pediatric Appendicitis Score i skala Alvarado, stanowią użyteczne narzędzia wspomagające,

szczególnie w warunkach oddziałów ratunkowych, jednak ich ograniczona czułość wskazuje na potrzebę uzupełniania ich o badania obrazowe. Ponadto algorytmy uczenia maszynowego, a zwłaszcza modele drzew decyzyjnych, mogą znacząco poprawić trafność predykcji diagnozy, sposobu leczenia i ciężkości choroby, oferując potencjał do tworzenia systemów wspomagania decyzji klinicznych w czasie rzeczywistym.

Z naukowego punktu widzenia przegląd literatury potwierdza, że ostre zapalenie wyrostka robaczkowego u dzieci pozostaje wyzwaniem diagnostycznym, a badania nad nowymi biomarkerami (np. wskaźnik SII) oraz zaawansowanymi metodami analizy danych przesuwały granice dotychczasowej wiedzy. Wykazano, że modele nieliniowe istotnie przewyższają tradycyjną regresję logistyczną w zadaniach klasyfikacji. Opracowanie i walidacja zewnętrzna indeksów takich jak BIDIAP [Arredondo Montero J., Pascual C., i inni, 2023] stanowią krok w kierunku medycyny precyzyjnej, choć konieczne są dalsze badania prospektywne na większych grupach badawczych.

Należy nadmienić, że metody uczenia maszynowego nie zastąpią oceny klinicznej, ale mogą stać się jej cennym, obiektywnym uzupełnieniem, pod warunkiem starannej walidacji i uwzględnienia ograniczeń, takich jak operatorozależność ultrasonografii czy niedobór danych dla przypadków powikłań.

ORCID iD

Vanessa Stabla: <https://orcid.org/0009-0004-0101-0355>

Małgorzata Zdrodowska: <https://orcid.org/0000-0003-4383-5713>

Literatura

1. Arredondo Montero J., Pascual C., Antona G., Ros Briones R., López-Andrés N., Martín-Calvo N. (2023), *The BIDIAP index: a clinical, analytical and ultrasonographic score for the diagnosis of acute appendicitis in children*, *Pediatric Surgery International* 39, art. 199, <https://doi.org/10.1007/s00383-023-05463-5>.
2. Dadeh A.A., Puitong K. (2021), *Predictive factors to diagnose appendicitis in children in the emergency department*, *Open Access Emergency Medicine* 13, s. 363-372, <https://doi.org/10.2147/OAEM.S323960>.
3. Marcinkevičs R., Reis Wolfertstetter P., Wellmann S., Knorr C., Vogt J.E. (2021), *Using machine learning to predict the diagnosis, management and severity of pediatric appendicitis*, *Frontiers in Pediatrics* 9, art. 662183, <https://doi.org/10.3389/fped.2021.662183>.

4. Marcinkevičs R., Reis Wolfertstetter P., Klimiene U., Ozkan E., Chin-Cheong K., Paschke A., Zerres J., Denzinger M., Niederberger D., Wellmann S., Knorr C., Vogt J.E. (2023), *Regensburg Pediatric Appendicitis Dataset*, Zenodo, <https://doi.org/10.5281/zenodo.7669442>.
5. Marcinkevičs R., Reis Wolfertstetter P. i in. (2023), *Regensburg Pediatric Appendicitis [Dataset]*, UCI Machine Learning Repository.
6. Parveen K., Avabratha Kadke S., Shetty K. (2017), *Pediatric appendicitis score in the diagnosis of childhood appendicitis: a validation study*, International Journal of Contemporary Pediatrics 4(6), s. 2196-2201, <https://doi.org/10.18203/2349-3291.ijcp20174756>.
7. Turiać A., Zdrodowska M. (2022), *Data Mining Approach in Diagnosis and Treatment of Chronic Kidney Disease*, Acta Mechanica et Automatica 16, s. 180-188, <https://doi.org/10.2478/ama-2022-0022>.
8. Zdrodowska M. (2019), *Attribute Selection for Stroke Prediction*, Acta Mechanica et Automatica 13(3), s. 200-204, <https://doi.org/10.2478/ama-2019-0026>.
9. Zdrodowska M., Dardzińska-Głębocka A. (2022), *Classification and action rules in identification and self-care assessment problems*, Technology and Health Care 30(1), s. 257-269, <https://doi.org/10.3233/THC-219008>.
10. Zdrodowska M., Kasperczuk A., Dardzińska-Głębocka A. (2023), *Selected feature selection methods for classifying patients with Hepatitis C*, Procedia Computer Science 225, s. 3710-3717, <https://doi.org/10.1016/j.procs.2023.10.366>.

Diagnosis of Appendicitis in Children Based on Clinical and Laboratory Findings

Abstract

Acute appendicitis is the most common surgical condition in children, accounting for one-third of hospital admissions for abdominal pain in this age group [Dadeh, Puitong, 2021; Marcinkevičs et al., 2023]. Despite the availability of diagnostic scales and advances in imaging, diagnosis in children remains a significant clinical challenge due to the atypical presentation of symptoms and limited ability to communicate with young patients [Arredondo Montero et al., 2023; Dadeh, Puitong, 2021]. Identifying reliable predictors and optimising management strategies, including differentiating between surgical and conservative treatment, may help reduce the incidence of complications and improve treatment efficacy [Marcinkevičs et al., 2021; Marcinkevičs et al., 2023].

The aim of the study was to analyse the relationship between the incidence of acute appendicitis in children and selected clinical, laboratory and demographic varia-

bles. The analysis took into account clinical symptoms, inflammatory parameters, ultrasound findings and scores on selected diagnostic scales. The analysis was conducted on the “Regensburg Pediatric Appendicitis” dataset available in the Zenodo repository. The data included demographic information (age, sex, BMI), clinical symptoms (including migration of pain, tenderness in the right lower quadrant, nausea, vomiting), laboratory test results (leukocytosis, CRP, neutrophilia) and ultrasound parameters (appendix diameter, presence of free fluid). Diagnostic scores (Alvarado Score, Paediatric Appendicitis Score) and outcome variables, such as diagnosis and treatment method, were also included.

Key words

acute appendicitis, paediatrics, diagnosis, Alvarado Score, Pediatric Appendicitis Score