

**SVEUČILIŠTE U SPLITU
MEDICINSKI FAKULTET**

IVAN MALEŠ, dr. med.

**MODELI STROJNOG UČENJA U DIJAGNOSTICI AKUTNOG
APENDICITISA U DJECE I ODRASLIH**

DOKTORSKI RAD

SPLIT, 2025.

Ovaj rad izrađen je na Katedri za patofiziologiju Medicinskog fakulteta Sveučilišta u Splitu.

Voditelj rada:

prof. dr. sc. Joško Božić, dr. med.

ZAHVALA

„Alo, oce li to? Kad ces kuci“

„reka san ti ja...“

„ta ta“

Prvo se želim zahvaliti svima koji su na bilo koji način sudjelovali i pomogli u izradi ovog rada te na dobivenim savjetima. Zahvaljujem se i prof. Pogoreliću koji je na samom početku nesebično podijelio svoje podatke čime je znatno ubrzana i olakšana provedba rane faze istraživanja.

Posebno hvala mentoru i prijatelju Jošku Božiću na podršci od samog početka ovog studija. Iako smo se na ovaj put upustili s drugim planovima, svojim povjerenjem, savjetima i ugađanjem raznim idejama pružio mi je slobodu samostalnog razvoja i ostavio trajan trag u mom profesionalnom i osobnom razvoju. Time je pokazao da mentorstvo nije samo davanje odgovora ili gotovih rješenja, nego strpljenje za dovoljan broj pitanja.

Hvala mojoj prvoj obitelji – ocu, majci i sestri, na dugogodišnjoj podršci.

Ovaj rad posvećujem Andrei i Lauri, bez kojih bi ovo životno poglavlje završilo i ranije, ali uz nemjerljivo manje radosti. Andrei hvala na strpljenju u trenucima kada je ovaj rad imao prednost nad svime ostalim, posebno u razdobljima kada je napredak bio spor ili teško uočljiv, čak i iz moje perspektive. Lauri koja je svojim noćnim buđenjima i ranojutarnjim osmjesima često davala do znanja da će ovaj rad pričekati. A oboma što su me podsjećale da su najvažnije stvari uvijek tu, često odmah iza ekrana.

Not all those who wander are lost

SADRŽAJ

1. UVOD.....	2
1.1. Crvuljak (appendix vermiformis)	2
1.1.1. Povijesni podaci	2
1.1.2. Anatomija i embriologija	3
1.1.3. Funkcija crvuljka	6
1.1.4. Neupalne bolesti crvuljka	7
1.1.5. Akutni apendicitis	12
1.1.5.1. Epidemiologija.....	13
1.1.5.2. Etiologija, patofiziologija i klasifikacija apendicitisa.....	14
1.1.5.3. Dijagnoza apendicitisa	15
1.1.5.4. Diferencijalna dijagnoza	21
1.1.6. Liječenje.....	22
1.1.6.1. Konzervativni pristup.....	22
1.1.6.2. Kirurški pristup	23
1.2. Strojno učenje	26
1.2.1. Modeli strojnog učenja	28
1.2.1.1. Logistička regresija.....	29
1.2.2. Validacija modela strojnog učenja.....	35
1.2.3. Objašnjivost modela strojnog učenja	37
1.2.4. Strojno učenje u medicini i abdominalnoj kirurgiji	37
1.2.5. Modeli strojnog učenja u akutnom apendicitisu	38
2. CILJEVI I HIPOTEZE ISTRAŽIVANJA	41
3. ISPITANICI I POSTUPCI	43
3.1. Ispitanici.....	44
3.2. Postupci.....	44
3.2.1. Izrada modela i optimizacija.....	46
3.3. Statistička analiza podataka	49
3.4. Objašnjivost modela.....	50
4. REZULTATI.....	51
4.1. Rezultati istraživanja u pedijatrijskoj populaciji.....	52
4.2. Rezultati istraživanja u odrasloj populaciji.....	61
5. RASPRAVA	74
6. ZAKLJUČCI.....	86

7. SAŽETAK.....	89
8. SUMMARY.....	92
9. LITERATURA	95
10. ŽIVOTOPIS.....	110

POPIS OZNAKA I KRATICA

AI – umjetna inteligencija (engl. *artificial intelligence*)

ANN – umjetna neuronska mreža (engl. *artificial neural network*)

AUC –površina ispod ROC krivulje (engl. *area under the curve*)

BMI – indeks tjelesne mase (engl. *body mass indeks*)

CT – kompjuterizirana tomografija (engl. *computed tomography*)

CVS – engl. *critical view of safety*

DALY – godine života prilagođene invaliditetu (engl. *disability-adjusted life years*)

DL – duboko učenje (engl. *deep learning*)

DT – stablo odluke (engl. *decision tree*)

GALT – limfno tkivo sluznice crijeva (engl. *gut-associated lymphoid tissue*)

IBD – upalna bolest crijeva (engl. *inflammatory bowel disease*)

LR – logistička regresija (engl. *logistic regression*)

MCHC – prosječna koncentracija hemoglobina u eritrocitu (engl. *mean corpuscular hemoglobin concentration*)

MiNEN – miješane neuroendokrine-neneuroendokrine neoplazme (engl. *mixed neuroendocrine-nonneuroendocrine neoplasm*)

ML – strojno učenje (engl. *machine learning*)

MPV – prosječan volumen trombocita (engl. *mean platelet volume*)

NEC – neuroendokrini karcinom (engl. *neuroendocrine carcinoma*)

NEN – neuroendokrini neoplazma (engl. *neuroendocrine neoplasm*)

NET – neuroendokrini tumor (engl. *neuroendocrine tumour*)

NLR – omjer neutrofila i limfocita (engl. *neutrophil to lymphocyte ratio*)

NOTES – endoskopska kirurgija kroz prirodne tjelesne otvore (engl. *natural orifice transluminal endoscopic surgery*)

TLR – omjer trombocita i limfocita (engl. *thrombocyte to lymphocyte ratio*)

RDW – raspodjela eritrocita po volumenu (engl. *red blood cell distribution width*)

RF – slučajne šume (engl. *random forest*)

ROC – krivulja međuodnosa osjetljivosti i specifičnosti (engl. *Receiver Operating Characteristics*)

SILS – laparoskopska kirurgija s jednom incizijom (engl. *Single incision laparoscopic surgery*)

SMOTE – engl. *Synthetic minority oversampling technique*

XGB – engl. *eXtreme Gradient Boosting*

WHO – Svjetska zdravstvena organizacija (engl. *World Health Organisation*)

YLD – godine života izgubljene invalidnošću (engl. *years lost due to disability*)

YLL – godine života izgubljene prijevremenom smrću (engl. *years of life lost*)

1. UVOD

1.1. Crvuljak (appendix vermiformis)

1.1.1. Povijesni podaci

Crvuljak, službenog anatomskog naziva *appendix vermiformis*, divertikularna je struktura koja izvire iz cekuma (1, 2). Povijesno gledano, crvuljak je vjerojatno opisan još u doba starih Egipćana oko 3000. godina pr. Kr. Tijekom procesa mumifikacije, abdominalne strukture su bile ekstrahirane i sačuvane u kanopskim vazama. Na tim posudama nalazili su se natpisi koji su opisivali njihov sadržaj, a za ono što je vjerojatno bio crvuljak koristili su izraz „intestinalni crv“ (2).

Renesansni polihistor Leonardo Da Vinci prvi je dokumentirao postojanje crvuljka jasnim prikazom u svojim crtežima 1492. godine. Potom je Berengario Da Carpi 1524. godine prvi opisao organ kao „slijep, sakuliran i jednook“ (3, 4). Izvori se razlikuju u pogledu toga tko je prvi dao naziv crvuljku. Neki navode da je Vido Vidius (latinizirano za Guido Guidi) opisao crvoliku strukturu kao vermiformni apendiks još 1530. godine. Drugi navode Andreasa Vesaliusa kako ga je nazivao slijepim crijevom/cekumom u svome djelu *De fabrica humani corporis libri septem* izdanog 1543. godine („*Hoc intestinum mihi caecum nuncupatur...*“) i da je prvi usporedio apendiks s crvom („*vermis in mondo convolutus*“) u drugom izdanju 1555. godine, a tek onda Eustahije u posthumno tiskanim anatomskim prikazima u 1714. godini prvi koristi izraz crvoliki („*vermiformis*“) uz ovu strukturu i navodi da naziv „slijepo“ jedni koriste za cekum, a drugi za crvuljak. Naposljetku, navodi se da je Verheye 1710. smislio naziv „*appendix vermiformis*“ na marginama svojih bilješki (3-5). Mnogi su opisivali slučajeve apendicitisa, ali sam naziv *appendicitis* dao je Fitz (5).

Mestevier je prvi objavio 1759. godine o operacijskom liječenju drenaže apscesne kolekcije zbog bolesti apendiksa (6). Najpoznatiji rani zapis o apendektomiji pripisuje se Amyandu koji je 1735. godine operirao jedanaestogodišnjeg dječaka zbog desnostrane preponske kile s enterokutanom fistulom. Međutim, zapis Williama Cookesleya koji je objavljen 1742., opisuje operaciju uklještene preponske kile iz 1731. U tom zapisu se navodi da je Cookesley za vrijeme operacije ekscidirao nekrotično tanko crijevo nakon čega se pacijent oporavio. Kada je pacijent umro 31 godinu kasnije, John Symons napravio obdukciju na kojoj je vidio intaktno tanko crijevo, ali je nedostajao dio cekuma i crvuljak. Stoga je Cookesley prvi, doduše nenamjerno, koji je izveo apendektomiju (6). Prva dokumentirana apendektomija kroz abdominalnu inciziju izvedena je 1880. godine kada je Lawson Tait operirao sedamnaestogodišnju djevojku zbog toga što je smatrao da je razvila generalizirani peritonitis

(6). Charles H. McBurney je 1889. godine u svom radu zagovarao rano operacijsko liječenje i opisao mjesto palpacijske bolnosti koje je kasnije nazvano McBurnyjeva točka, dok je u radu objavljenom 1894. godine opisao kosu abdominalnu inciziju, koja je također dobila ime po njemu (7, 8).

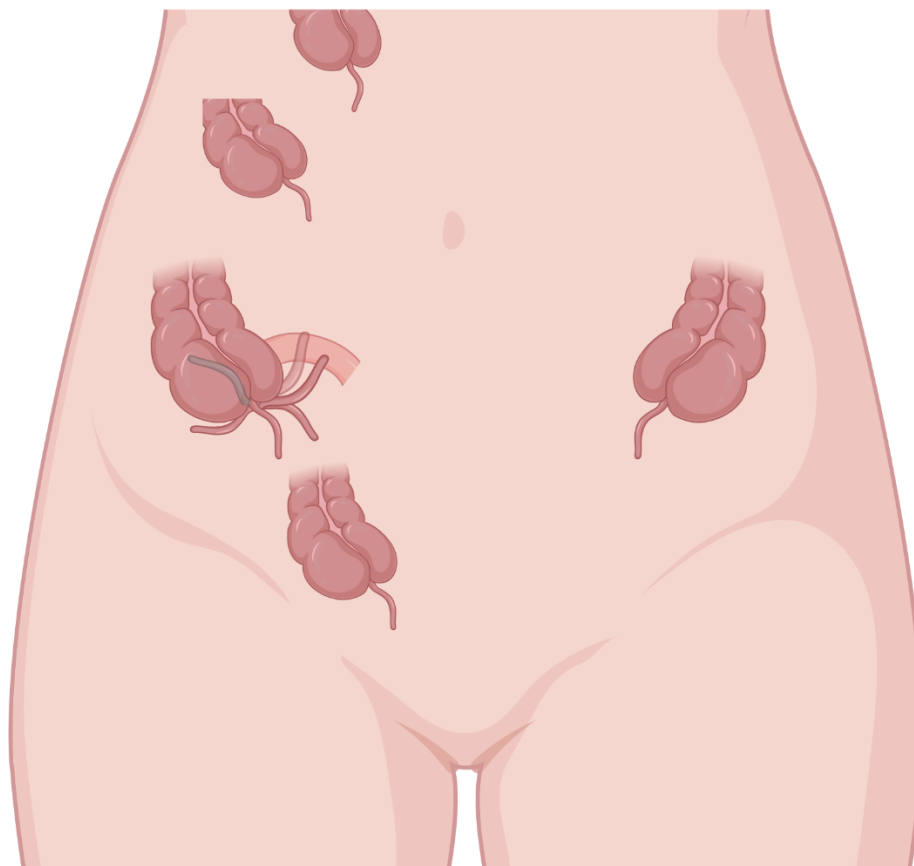
Početak 20. stoljeća počelo se razmatrati i o odgođenom operacijskom liječenju. Prvi je to zabilježio Ochsner 1902., a zatim je dodatno modificirano od strane Sherrena i nazvano Ochsner-Sherren protokol (9). Uporabu antibiotske terapije u liječenju akutnog apendicitisa zagovarao je Coldrey, koji je 1956. godine objavio rad u kojem je usporedio operacijsko liječenje s konzervativnim liječenjem antibioticima (10).

Razvojem ginekološke dijagnostičke laparoskopije, Semm je 1983. godine uspješno izveo prvu laparoskopsku apendektomiju (11). Daljnjim razvojem tehnologije i kirurških tehnika uspješno su izvedene laparoskopske apendektomije kroz jedno mjesto incizije (engl. *single incision laparoscopic surgery*, SILS) te apendektomije u sklopu endoskopskog kirurškog pristupa kroz prirodne tjelesne otvore kao što su transgastrična, transvaginalna, transvezikalna pa čak i transkolonična (kolonoskopska) apendektomija (12-16). Interesantno je spomenuti i slučajeve samostalno izvedenih apendektomija, u kojima je kirurg operirao samoga sebe. Najpoznatiji je slučaj ruskog kirurga koji je tijekom ekspedicije na Antarktiku 1961. godine bio primoran izvesti operaciju na sebi, nakon što nije došlo do poboljšanja na antibiotsku terapiju. No, prvi pokušaj izveo je Bertram F. Aldena 1912. godine, dok je prvu zabilježenu uspješnu samostalno izvedenu apendektomiju napravio je Evan O'Neill Kane 1921. godine (6).

1.1.2. Anatomija i embriologija

Crvuljak embriološki predstavlja ekstenziju cekuma. To je slijepo završena tubularna struktura varijabilnog promjera, najčešće od 0,5 do 0,8 cm, kao i varijabilne duljine, u prosjeku oko 8 cm. Izlazište mu je iz posteromedijalnog dijela cekuma, oko 1 do 3 cm ispod ileocekalnog spoja, na mjestu gdje se tenije debelog crijeva spajaju u longitudinalni mišićni sloj crvuljka (17, 18). Iako je položaj baze crvuljka u odnosu na cekum konstantan, položaj vrška crvuljka izrazito je varijabilan (17). Crvuljak je vezan za ileum preko mezoapendiksa (19). Najčešće opisani položaj crvuljka je retrocekalni (oko 32%), a zatim pelvični (oko 28%). Od ostalih opisanih položaja spominju se subcekalni, ilealni (koji se dijeli na pre- i postilealni), paracekalni, anterocekalni i hepaticični te ostali. Zabilježena je i razlika između položaja inflamiranog i inocentnog crvuljka, pa je tako inflamirani češće bio položen retrocekalno, a

inocentan pelvično (18). Osim toga, položaj cekuma također može biti varijabilan što za posljedicu ima i drugačiji položaj crvuljka (20). Ovi različiti mogući položaji su prikazani na **Slici 1.**



Slika 1. Anatomske varijacije u položaju cekuma i crvuljka. Kreirano koristeći *BioRender*.

Razvoj probavnog sustava počinje oko četvrtog tjedna gestacije razvojem primitivnog crijeva. Ono je sačinjeno od tube endoderma koji je okružen glatkim mišićjem mezoderma. U konačnom razvoju sudjeluje i ektoderm iz kojeg se razvija enterični živčani sustav. Primitivno crijevo se dijeli na prednje, srednje i stražnje što je i osnova podjele vaskularne opskrbe probavnog sustava (21). Daljnji razvoj se može podijeliti u tri stadija. U prvom srednje crijevo ulazi u ekstraembrijonalni celom i postepeno se elongira u petlju oblika „V“ koja se širi naprijed prema umbilikusu. U toj fiziološkoj hernijaciji smješta se veći dio probavnog sustava do trećeg tromjesečja trudnoće kada se događa rotacija oko osi gornje mezenterične arterije (lat. *arteria mesenterica superior*, AMS). U drugom stadiju dolazi do daljnje rotacije, sve do ukupno 270 stupnjeva, a tada se i duodenum fiksira posteriorno od AMS. Treći stadij obilježen je fiksacijom kolona i pripadajućeg mezenterija u trbušnu šupljinu. Uzlazno i silazno debelo crijevo s mezenterijem se fiksiraju u trbušnu šupljinu čime se formira Toldtova fascija koja je granica između mezenterija i retroperitoneuma. Poprečno debelo crijevo se spoji s velikim

omentumom čime se formira omentalna burza. Cekalni pupoljak je posljednji dio debelog crijeva koje se smjesti u trbušnu šupljinu. Inicijalno smješten u gornjem desnom kvadrantu, s vremenom se izduži i zauzme svoje mjesto u donjem desnom kvadrantu (22). Razvoj crvuljka počinje tijekom osmog tjedna gestacije. Inicijalno položen na samom vršku cekuma, razvojem hastrua položaj mu se mijenja prema ileocekalnoj valvuli. Regionalni se limfni čvorovi počinju razvijati sedmog mjeseca gestacije i nastavljaju se proliferirati do puberteta kada počinju regredirati (21).

Strukture koje se embriološki razvijaju iz srednjeg crijeva vaskularnu opskrbu dobivaju preko AMS. Postoje tri grane za debelo crijevo koje mogu imati podrijetlo iz AMS: srednja količna, desna količna i ileokolična arterija. Ileokolična arterija se dalje dijeli u dvije grane, od kojih donja grana daje apendikularnu arteriju za arterijsku opskrbu crvuljka (22). Venska drenaža odgovara arterijskog opskrbi pa se tako strukture embriološkog srednjeg crijeva dreniraju preko gornje mezenterične vene (lat. *vena mesenterica superior*, VMS). Postoje četiri grupe limfne drenaže opisane za debelo crijevo: epiploične, parakolične, intermedijalne i glavne (22). Inervacija struktura embriološki nastalih iz srednjeg crijeva dolazi od simpatičkih vlakana od razine Th6 do Th12, a postganglijska vlakna su položena uz AMS. Nadalje, parasimpatička inervacija dolazi preko vagalnog živca i celijačnog pleksusa (22).

Stijenka crvuljka sastoji se od četiri glavna sloja: sluznice, podsluznice, mišićnog sloja i seroze, slično strukturi debelog crijeva. Mukoza se sastoji od cilindričnog epitela s enterocitima i vrčastim stanicama, lamine proprije i *muscularis mucosae*. Osim makrofaga, u lamini propriji se nalazi obilje plazma stanica koje proizvode imunoglobulin (Ig) A ili IgG. Lieberkühnove kripte su također prisutne, slično kao u debelom crijevu. Panethove stanice, koje se uobičajeno nalaze u tankom crijevu, smještene su na dnu ovih kripti, a njihova glavna funkcija je proizvodnja antimikrobnih peptida. Podsluznica se sastoji od vezivnog tkiva i karakterizirana je prisutnošću brojnih limfoidnih folikula koji se protežu iz submukoze u laminu propriju. Iako njihova prisutnost ili ekvivalentna struktura nije uočena u zdravom debelom crijevu, usporedivi su s Peyerovim pločama u tankom crijevu. Zona plašta ovog limfoidnog tkiva, lokalizirana najbliže lumenu, sadrži gusto zbijene B limfocite i malo T limfocita. Tamna zona unutar jasnog germinativnog centra nalazi se najdalje od lumena i sadrži makrofage i centroblaste, proliferirajuće B stanice koje monoklonalnom ekspanzijom stvaraju folikul. Centroblasti se diferenciraju u centrocite, koji zajedno s folikularnim dendritičkim stanicama čine svijetlu zonu. Između kupolastog epitela i limfoidnih folikula nalazi se još jedno područje imunskih stanica: zona mješovitih stanica, koja sadrži makrofage i limfocite, uključujući B i T stanice. Na dnu limfoidnih folikula nalaze se T-stanična područja koja sadrže

makrofage i T stanice, s osam puta više CD4⁺ nego CD8⁺ T stanica. Mišićni sloj sastoji se od glatkih mišića odgovornih za peristaltičke pokrete, dok vanjski sloj seroze pruža zaštitni omotač (23). Limfno tkivo je obilnije u crvuljku nego u debelom crijevu, što ide u prilog imunološke funkcije crvuljka (24).

1.1.3. Funkcija crvuljka

Dugo vremena funkcija crvuljka bila je nepoznata. Primjerice, Berengario da Carpi smatrao je da je uloga crvuljka pomoćna pumpa koja bi olakšala peristaltiku isisavanjem sadržaja iz ileuma i izbacivanjem u debelo crijevo. Za razliku od njega, Darwin je smatrao da je crvuljak rudimentarna, zakržljala struktura (4, 25). S tehnološkim napretkom u mikroskopiji, otkriveno je da stijenka crvuljka sadržava znatnu količinu limfnog tkiva sluznice crijeva (engl. *gut-associated lymphoid tissue*, GALT) (25). Imunoglobulin A (IgA) i mucin, proteini koji se nalaze u većim količinama u GALT-u, potpomažu u rastu bakterijskih kolonija u vidu biofilma (26). Bollinger i sur. 2007. godine objavili su važan rad za razumijevanje funkcije crvuljka. U njemu su predložili da je crvuljak utočište za komenzalne mikroorganizme. Ti mikroorganizmi stvaraju biofilme koji se otpuštaju u debelo crijevo, što bi imalo ulogu u redovitoj regeneraciji crijevne flore, osobito nakon crijevnih bolesti uzrokovanih patogenim mikroorganizmima čime bi funkcija crvuljka bila skladište komenzalnih bakterija za rekolonizaciju debelog crijeva (27). U skladu s tim, pokazano je da postoji veći rizik za rekurentne infekcije, kao i da utječe na težinu *C. difficile* infekcije te da apendektomirani bolesnici imaju lošije ishode (28-30).

Funkcija crvuljka može se očitovati i u sudjelovanju u povezanosti s upalnim bolestima crijeva (engl. *inflammatory bowel disease*, IBD). Česta je pojava upalnog zbivanja koje zahvaća ušće crvuljka kod bolesnika s IBD-om čak i u slučajevima kada je upala lokalizirana samo na lijevi kolon ili u rektumu. Također, kod nekih bolesnika pojava upalnog zbivanja oko ušća crvuljka prethodi ulceroznom kolitisu, sugerirajući da ima pokretačku ulogu ili ulogu u razvoju bolesti te da se može koristiti kao rani endoskopski marker za razvoj bolesti kod nekih bolesnika (23, 31). Provedena velika kohortna istraživanja pokazala su povezanost apendektomije s IBD-om. Većina studija slučajeva i kontrola koja su evaluirala ulogu apendektomije u prevenciji razvoja IBD-a pokazuju značajnu inverznu povezanost apendektomije i razvoja UC-a (32). Kohortno istraživanje koje su proveli Andersson i sur. na 400 000 ispitanika pokazalo je da kod osoba koje su apendektomirane zbog akutnog apendicitisa ili mezenterijalnog limfadenitisa postoji manja stopa incidencije UC-a u usporedbi

s kontrolnom skupinom ako su imale manje od 20 godina u vrijeme apendektomije (33). Također, Myrelid i sur. pokazali su da je apendektomija prije razvoja UC-a povezana s manje agresivnim oblikom bolesti tako što su zabilježili manji rizik od kolektomija i hospitalizacija vezanih uz UC. Međutim, pronašli su i lošiju dugoročnu prognozu kod bolesnika s UC-om ako su bili apendektomirani nakon postavljene dijagnoze UC-a (34). Frisch i sur., u svom istraživanju slučajeva i kontrola s 30 632 sudionika, ispitivali su populaciju stariju od 20 godina u vrijeme apendektomije te nisu pronašli inverznu poveznicu s apendektomijom i rizikom za razvoj ulceroznog kolitisa. Povećan rizik od UC-a, a osobito CD-a, u ranom razdoblju nakon apendektomije povezuju s diferencijalno dijagnostičkim problemom kod pacijenata koji se prezentiraju prvi put s bolovima u trbuhu. (35). Za razliku od njih, Andersson i sur. pronašli su i povećan rizik od razvoja CD-a i do 20 godina nakon apendektomije stoga ne smatraju točnim hipotezu incipijentnog CD-a kao uzroka bolova u trbuhu koji su doveli do apendektomije (36). Međutim, na još većim kohortnim istraživanjima s duljim periodom praćenja, kao i na provedenoj meta-analizi, potvrđen je povećan rizik razvoja CD-a nakon apendektomije, kao i nestanak povećanog rizika u većini podgrupa nakon 5 godina, što bi sugeriralo da dijagnostička pristranost objašnjava povećan rizik (37, 38). Nadalje, veliko nacionalno kohortno istraživanje na tajvanskoj populaciji starijoj od 20 godina pokazalo je veću incidenciju i rizik od CD-a i UC-a kod osoba koje su apendektomirane u usporedbi s neapendektomiranom kohortom. Svoje rezultate povezuju s rasnim varijacijama i disreguliranom crijevnom mikrobiotom zbog okolišnih utjecaja (39). Veliko kohortno kanadsko istraživanje iz 2022. godine potvrđuje prethodna saznanja o povećanom riziku CD-a nakon apendektomije što također objašnjavaju dijagnostičkom pristranošću, dok potvrđuju protektivnu poveznicu apendektomije i razvoja UC-a (40). Provedena istraživanja o utjecaju apendektomije na tijek već dijagnosticiranog UC-a dala su dijelom oprečne rezultate, iako su većinom izvijestila o pozitivnom učinku na tijek bolesti (32). Preliminarni rezultati randomiziranog kontroliranog istraživanja kojeg provode Visser i sur. pokazuju da apendektomija može utjecati na stopu ozbiljnih nepoželjnih događaja povezanih s kolitisom kao i na stopu relapsa (41).

1.1.4. Neupalne bolesti crvuljka

Bolesti crvuljka mogu se klasificirati na upalne bolesti, stanja vezana za kongenitalne anomalije i druge neupalne bolesti te tumore crvuljka (24).

1.1.4.1. Kongenitalne anomalije

Duplikature crvuljka izrazito su rijetke anomalije, s manje od 200 dokumentiranih slučajeva u svijetu. Pravu incidenciju teško je procijeniti jer, osim što je riječ o rijetkoj anomaliji, veliki broj slučajeva neće se niti identificirati. Usprkos tome, riječ je o bitnom medicinsko-pravnom entitetu te bi, uzimajući u obzir frekvenciju apendektomija, svi kirurzi trebali biti upoznati s mogućnošću dvostrukog crvuljka (42). One se mogu podijeliti prema modificiranoj Cave-Wallbridge klasifikaciji, a predložena je i potreba za modifikacijama iste pa su tako Ngulube i sur. dali do sada najopširniju klasifikaciju (43, 44). Najčešće opisani oblik duplikature je oblik B (42). Također, postoje i opisani još rjeđi slučajevi triplikatura crvuljka, kao i ageneze (45, 46). Potonji se mogu, zajedno s agenezom cekuma, svrstati po Collinsovoj klasifikaciji u pet tipova (47). Međutim, hipoplazija ili intramuralna lokacija mogu dovesti do pogrešnog prikaza pravih stopa incidencije (48). U nekim slučajevima uzrok nije kongenitalan, već može biti sekundaran zbog intrauterinog vaskularnog incidenta, autoamputacije ili uzimanja lijekova tijekom trudnoće (npr. talidomid) (49).

1.1.4.2. Neupalne bolesti crvuljka

Divertikuloza crvuljka je također rijetko stanje s opisanom incidencijom od 0,014% do 3,7%. Ono može biti kongenitalno ili stečeno. Stečeni pseudodivertikli često su lokalizirani u distalnoj trećini crvuljka na mezenterijalnoj strani, dok su kongenitalni pravi divertikli i lokalizirani su na antimezenterijalnoj strani crvuljaka. Može se klasificirati u 4 morfološka tipa: akutni divertikulitis, akutni apendicitis s akutnim divertikulitisom, akutni apendicitis s divertiklom i crvuljak s divertiklom te u podskupine ovisno o postojanju perforacije. U izrazito rijetkim slučajevima (0,004%-2,1%) može biti uzrok akutnog apendicitisa. Povezana je s višim rizikom od neoplazmi, posebno karcinoidom i mucinoznim adenomima. Diferencijalno dijagnostički, ova dijagnoza dolazi u obzir kod sumnje na akutni apendicitis, a opisane su i više stope perforacije u odnosu na apendicitis (50, 51).

Endometrioza je bolest karakterizirana prisutnošću stanica sličnih endometriju van endo- i miometrija. Opisana je kao rijetko stanje i bolesnici mogu biti asimptomatski, ali se također mogu prezentirati nizom raznih gastrointestinalnih smetnji kao što su bolovi u trbuhu, krvarenje iz probavnog sustava ili pak kliničkom slikom akutnog apendicitisa (52).

Fibrozna obliteracija lumena crvuljka, još zvana i neurogena hiperplazija, je tipično incidentalan nalaz nakon apendektomije. Obično zahvaća vrh samog crvuljka, prezentirajući se kao blijedo do bijelo obliteriranje lumena. Histološki postoje tri varijante: centralna, koja je

i najčešća, submukozna i intramukozna varijanta. Iako može nalikovati nekim drugim benignim tumorima probavnog sustava, ovo je benigna lezija koja ne zahtijeva specifično liječenje osim apendektomije (53, 54).

Vaskulitis crvuljka ima učestalost od 0,3% do 1,9%, češće kod mladih žena. Može biti dio sustavnog vaskulitisa ili izolirana pojava. Može biti primarni vaskulitis ili sekundarno zbivanje uslijed akutnog apendicitisa. Nodozni poliarteritis najčešći je sustavni vaskulitis koji zahvaća slijepo crijevo. U nekim slučajevima vaskulitis crvuljka može prethoditi razvoju sustavnog nodoznog poliarteritisa. Ostali povezani poremećaji uključuju reumatoidni vaskulitis, sistemski lupus i ANCA-pozitivni mikroskopski poliarteritis (55).

Torzija crvuljka može se pojaviti kao primarno ili sekundarno zbivanje, pri čemu je većina slučajeva u pedijatrijskoj populaciji primarna. Simptomi su vrlo slični apendicitisu i izrazito je teško ovo stanje dijagnosticirati preoperativno, stoga se dijagnoza najčešće postavi tek intraoperativno (56).

Intususcepcija crvuljka je još jedno rijetko stanje čija klinička slika varira od akutne i progresivne, koja može imitirati akutnu upalu crvuljka, pa sve do kronične i protrahirane s povremenim simptomima. Češće se javlja u odraslih nego u djece. Većina slučajeva se dijagnosticira za vrijeme operacije, ali se pokazao trend boljeg dijagnosticiranja nakon 2000. godine u odnosu na razdoblje prije. Neki od mogućih uzroka nastanka ovog stanja, osim upale, su endometrioza, mukocela, adenom, karcinoid i adenokarcinom. Klasifikacija se zasniva na položaju na kojem je crvuljak invaginiran i opisano je 5 tipova po McSwainu (57).

Opisani su i slučajevi stranih tijela u crvuljku kao i pojave parazita poput *Enterobiusa vermicularis* u lumenu crvuljka. Osim toga, crvuljak može biti sekundarno zahvaćen u sustavnim bolestima kao što su cistična fibroza, Crohnova bolest, ulcerozni kolitis i Hirschsprungova bolest te sekundarno upalno izmijenjen (48).

1.1.4.3. Tumori crvuljka

Najčešći tumori crvuljka su epitelnog i mezenhimalnog podrijetla. Hematolimfoidni tumori se rijetko pronalaze u crvuljku. Neuroendokrini tumori (engl. *neuroendocrine tumour*, NET) su česti te se najčešće otkrivaju incidentalno. Adenokarcinomi slični onima koji se nalaze u debelom crijevu i rektumu rijetki su u slijepom crijevu. Kada se ovi tumori razviju u slijepom crijevu, dijele slične morfološke i genetske karakteristike sa svojim pandanima u kolorektalnom području. Po posljednjoj klasifikaciji Svjetske zdravstvene organizacije (engl. *World Health Organisation*, WHO) iz 2019. godine, epitelni tumori crvuljka ugrubo su se

klasificirali kao seratne lezije i polipi, mucinozne neoplazme, adenokarcinomi i neuroendokrine neoplazme. Detaljni prikaz je vidljiv u **Tablici 1.** (58). Važno je napomenuti da izraz pseudomiksom peritoneuma nije histološki utemeljen pojam stoga se ne koristi u određivanju stadija bolesti (59).

Tablica 1. Klasifikacija epitelnih tumora crvuljka prema WHO-u iz 2019.

Epitelni tumori	Tip tumora
Benigne i prekursorske lezije	Hiperplastični polip
	Sesilna seratna lezija bez displazije
	Seratna displazija niskog gradusa
	Seratna displazija visokog gradusa
Mucinozne neoplazme	Mucinozna neoplazma crvuljka niskog gradusa (LAMN)
	Mucinozna neoplazma crvuljka visokog gradusa (HAMN)
Adenokarcinomi	Adenokarcinom, NOS
	Mucinozni adenokarcinom
	Adenokarcinom prstena pečatnjaka
	Nediferencirani karcinom, NOS
	Adenokarcinom vrčastih stanica
Neuroendokrine neoplazme	
	<i>Dobro diferencirane</i>
	Neuroendokrini tumor, NOS
	- Gradus 1
	- Gradus 2
	- Gradus 3
	Tumor L stanica
	Tumor koji luči peptid sličan glukagonu
	Tumor koji luči PP/PYY
	Karcinoid enterokromafinih stanica
	Karcinoid koji luči serotonin
	<i>Slabo diferencirane</i>
	Neuroendokrini karcinom, NOS
	Velikostanični
	Malih stanica
<i>Miješani tip</i>	Miješana neuroendokrino-neuroendokrino neoplazma (MiNEN)

Kratice: HAMN: mucinozna neoplazma crvuljka visokog gradusa (engl. *high-grade appendiceal mucinous neoplasm*); LAMN: mucinozna neoplazma crvuljka niskog gradusa (engl. *low-grade appendiceal mucinous neoplasm*); NOS: nije drugačije specificirano (engl. *not otherwise specified*); PP/PYY: pankreatički polipeptid/peptid YY; MiNEN: miješana neuroendokrino-neuroendokrino neoplazma (engl. *mixed neuroendocrine-non neuroendocrine neoplasm*)

1.1.4.3.1. Benigne i prekursorske lezije

Polipi i nazubljene lezije crvuljka su sluznični epitelni polipi koje karakterizira nazubljena (pilasta ili zvjezdasta) arhitektura lumena kriptе. Dijeļe se na hiperplastične polipe, sesilne nazubljene lezije bez displazije i nazubljene lezije s displazijom niskog ili visokog stupnja. Većina ih se otkrije slučajno nakon apendektomije iz drugih razloga. Veće tvorbe mogu potencijalno uzrokovati opstrukciju lumena crvuljka i dovesti do apendicitisa. Posjeduju i maligni potencijal zbog KRAS i BRAF mutacija. Ključno ih je razlikovati od LAMN-a zbog malignog potencijala i posljedičnih kliničkih implikacija (58).

1.1.4.3.2. Mucinozne neoplazme crvuljka

Mucinozne neoplazme crvuljka su promjene karakterizirane mucinoznom epitelnom proliferacijom s izvanstaničnim mucinom i ekspanzivnim, potiskujućim rubovima tumora. Makroskopski crvuljak ne mora odavati dojam značajnih promjena ili može biti proširen kao rezultat abnormalnog nakupljanja mucina, a može biti prisutna i perforacija s ekstruzijom mucina. One se dijeļe na mucinoznu novotvorinu crvuljka niskog i visokog stupnja, od kojih su visokog stupnja rjeđe. Prognoza izrazito ovisi o stadiju proširenosti bolesti u trenutku dijagnoze za LAMN, dok su podaci za HAMN oskudniji. U ranim stadijima LAMN ima izvrsnu prognozu, dok je kod peritonealne diseminacije prognoza varijabilna (58).

1.1.4.3.3. Adenokarcinomi

Adenokarcinom crvuljka je maligna glandularna/žljezdana neoplazma karakterizirana invazijom kroz slojeve. Klinička slika je varijabilna, pa se može otkriti u sklopu apendicitisa ili bolova u trbuhu. Bolesnici se mogu žaliti na promjene u crijevnim navikama ili krvarenje iz probavnog sustava, a kliničkim pregledom se može pronaći palpabilna tvorba. Podtipovi su adenokarcinom prstena pečatnjaka, mucinozni adenokarcinom i nediferencirani adenokarcinom. Bolesnici s mucinoznim adenokarcinomima imaju bolju prognozu u odnosu na nemucinozne, a histološki gradus se pokazao kao nezavisni prognostički faktor za petogodišnje preživljenje kod mucinoznih karcinoma (58).

Po definiciji, adenokarcinom prstena pečatnjaka je tumor koji se sastoji od mucinoznih stanica prstena pečatnjaka kao i varijabilnog broja endokrinih stanica i stanica koje nalikuju Panethovim. Najčešće je lociran u distalnom dijelu crvuljka. Da bi bio klasificiran kao adenokarcinom vrčastih stanica, tumor mora sadržavati barem komponentu klasičnog adenokarcinoma vrčastih stanica niskog gradusa. Prognoza ovisi o stadiju i gradusu. Većina bolesnika se dijagnosticira u stadiju 1 ili 2 kod niskog gradusa, dok se većina visokog gradusa

dijagnosticira u četvrtom stadiju. Najčešća sjela presadnica su peritoneum, omentum, trbušna stijenka i jajnici (58).

1.1.4.3.4. Neuroendokrine neoplazme crvuljka

Neuroendokrine neoplazme (engl. *neuroendocrine neoplasm*, NEN) crvuljka su epitelne neoplazme crvuljka s neuroendokrinom diferencijacijom, što uključuje i dobro diferencirane neuroendokrine tumore (engl. *neuroendocrine tumor*, NET) i slabo diferencirane neuroendokrine karcinome (engl. *neuroendocrine carcinoma*, NEC). Miješane neuroendokrine-neneuroendokrine neoplazme (engl. *mixed neuroendocrine-nonneuroendocrine neoplasm*, MiNEN) su miješane epitelne neoplazme s neuroendokrinom komponentom kombiniranom s neneuroendokrinom komponentom, a obje se mogu morfološki i imunohistokemijski prepoznati te svaka mora sudjelovati u najmanje 30% neoplazme (58). NET crvuljka klasificira se drugačije od NET-ova probavne cijevi koji se razvijaju iz embriološkog srednjeg crijeva (ileum i jejunum) zbog svoje učestalosti i posebnih obilježja. NET crvuljka ima udio od 5,7% svih NET-ova i 85% svih maligniteta crvuljka. Često se dijagnosticiraju u mlađoj životnoj dobi i imaju odličnu prognozu (60). Većina ih se dijagnosticira incidentalno nakon napravljene apendektomije koja je često dostatna kao zahvat. Ovisno o stadiju i visokorizičnim značajkama, kod određenog broja bolesnika dolazi u obzir desna hemikolektomija, a kod metastatske bolesti pristupa se onkološkom liječenju i eventualnoj kirurškoj citoredukciji. Slabo diferencirani neuroendokrini karcinomi se rijetko pojavljuju u crvuljku. Većina bolesnika s NEC-om probavnog sustava imaju metastatsku bolest u trenutku dijagnoze, te je liječenje onkološko. Dijagnoza MiNEN-a se postavlja izrazito rijetko, ali je crvuljak često primarno sjelo. Neuroendokrina komponentna je voditelj biološkog ponašanja, a najčešće je to slabo diferencirani neuroendokrini karcinom. Kirurško liječenje dolazi u obzir u potencijalno izlječivom ranom stadiju bolesti (61).

1.1.5. Akutni apendicitis

Akutna bol u trbuhu čini 7-10% svih prijema u hitnoj službi, a akutni apendicitis je najčešće intraabdominalno zbivanje u kojem je u većine bolesnika potrebno hitno kirurško liječenje (62, 63).

1.1.5.1. Epidemiologija

Analiza provedena od strane GBD suradničke skupine iz 2021. godine pokazala je da je globalna dobno-standardizirana stopa incidencije upale slijepog crijeva iznosila 214 (174–274) na 100 000, što je ekvivalentno 17 milijuna (13,8–21,6) novih slučajeva. Najviša stopa incidencije zabilježena je u visokorazvijenoj regiji Azije i Pacifika, s 364 (286–475) na 100 000, dok je najniža bila u zapadnoj podsaharskoj Africi, sa 81,4 (63,9–109) na 100 000. Globalne dobno-standardizirane stope smrtnosti, incidencije, godina života izgubljene prijevremenom smrću (engl. *years of life lost*, YLL), godina života izgubljenih invalidnošću (engl. *years lost due to disability*, YLD) i godina života prilagođenih invaliditetu (engl. *disability-adjusted life years*, DALY) povezane s upalom slijepog crijeva pokazale su kontinuirani pad između 1990. i 2021. godine, pri čemu je najveće smanjenje zabilježeno kod smrtnosti i YLL-ova. Najveći godišnji pad DALY-ova uočen je kod djece mlađe od 10 godina. Iako su stope smrtnosti smanjene u svim regijama, postojale su značajne regionalne razlike u vremenskim trendovima incidencije. Premda je globalna dobno-standardizirana stopa incidencije upale slijepog crijeva opadala između 1990. i 2021. godine, u gotovo polovici GBD regija zabilježen je porast incidencije od preko 10%. Stope smrtnosti varirale su od 1,01 (0,895–1,13) na 100 000 u središnjoj Latinskoj Americi do 0,054 (0,0464–0,0617) na 100 000 u visokorazvijenoj regiji Azije i Pacifika. Dobno-standardizirana stopa incidencije iznosila je 214 (174–274) na 100 000, što predstavlja smanjenje od 2,26% (0,693–3,82) u odnosu na 1990. godinu, kada je stopa iznosila 219 (95% UI 176–283) na 100 000, usprkos značajno većem broju novih slučajeva na globalnoj razini. Najviša stopa incidencije zabilježena je u drugom i trećem desetljeću života, globalno i u svih sedam GBD super-regija. Najveći godišnji pad incidencije primijećen je kod djece mlađe od 10 godina, neovisno o spolu. Premda je incidencija bila najviša u drugom i trećem desetljeću života, geografske varijacije pokazale su da se najveći teret bolesti razlikuje ovisno o dobnoj skupini u različitim regijama (64).

Na razini Hrvatske, 1990. godine zabilježene su 22 smrti (10 do 25.5 95% UI), a dobno standardizirana stopa smrtnosti je bila 0,407 na 100 000. Usporedno s time, u 2021. godini procijenjeno je da je bilo 9-10 smrtnih slučajeva. U tom razdoblju došlo je do pada u broju smrtnih slučajeva za 56% i pada u dobno-standardiziranoj stopi smrtnosti na 100 000 za 74% (64).

1.1.5.2. Etiologija, patofiziologija i klasifikacija apendicitisa

Opstrukcija lumena crvuljka smatra se patofiziološkim inicijatorom upale. Uslijed opstrukcije lumena dolazi do povišenja intraluminalnog tlaka i distenzije zbog nemogućnosti izlaska intraluminalnog sadržaja kojeg stvara mukoza. U daljnjem tijeku, povišeni intraluminalni tlak dovodi do okluzije kapilara i venula u stijenci crvuljka uz očuvanu arterijsku opskrbu, zbog čega dolazi do vaskularne kongestije što utječe na permeabilnost sluznice. Luminalno bakterijsko prerastanje, prvotno aerobnim mikroorganizmima, a zatim i anaerobnim, te translokacija tih mikroorganizama podržavaju daljnju upalu i edem. Uz to, kao i već opisan poremećaj protok krvi, dolazi do ishemije te u konačnici gangrene te perforacije. Inicijalno upala može biti zadržana okolnim tkivima, primjerice omentumom. Progresijom bolesti postupno dolazi do peritonitisa ili formiranja lokaliziranih apscesnih kolekcija (24, 65-67).

Međutim, ova teorija razvoja akutnog apendicitisa dovedena je u pitanje studijama koje su pokazale nisku učestalost opstruktivnih lezija u patološkim uzorcima i njihovu relativnu frekvenciju u patološki urednim uzorcima (22). Do opstrukcije može doći zbog limfoidne hiperplazije, fekolita, stranih tijela ili neoplazmi (65). Izrazito rijetko akutni apendicitis može nastati sekundarno zbog metastatske bolesti (68). U literaturi se opisuje i traumom uzrokovan apendicitis, ali je teško ustanoviti uzročno posljedičnu vezu (69). Opisani su razni virusni, bakterijski, gljivični i parazitarne uzročnici apendicitisa (**Tablica 2**) (70).

Tablica 2. Prikaz infektivnih uzroka koji zahvaćaju crvuljak

Bakterije	Virusi	Gljivice	Paraziti
<i>Salmonella sp</i>	Ospice	Mucormycosis	<i>Enterobius vermicularis</i>
<i>Shigella sp</i>	Adenovirus	Histoplazmoza	Shistosomes
<i>Yersinia sp</i>	Citomegalovirus		<i>Entamoeba histolytica</i>
<i>Acinomyces sp</i>	Epstein-Barr virus		<i>Balantidium coli</i>
<i>Campylobacter sp</i>			<i>Strongyloides stercoralis</i>
<i>Clostridium</i>			Toksoplazma
Mikobakterije			<i>Cryptosporidium</i>
<i>Rickettsia rickettsii</i>			Echinococcus
			<i>Trichuris sp</i>
			<i>Ascaris sp</i>

Prethodno opisana sekvenca koja u konačnici dovodi do perforacije nije neminovna već može doći i do spontane rezolucije stanja. Prema novijim studijama, na komplicirani i nekomplicirani apendicitis se može gledati kao dva odvojena entiteta. Prvi koji će progredirati do perforacije, i drugi koji se može spontano razriješiti. Tijekom COVID-19 pandemije

zabilježen je povećan omjer kompliciranih naspram nekompliciranih apendicitisa u odnosu na prepandemijsko razdoblje. Međutim, taj povećan omjer ne može se pripisati povećanom riziku za progresiju bolesti zbog ograničenog pristupa zdravstvenoj skrbi, već podaci sugeriraju na veći broj spontano razriješenih nekompliciranih apendicitisa. Na taj zaključak ukazuje pad u broju nekompliciranih apendicitisa, dok je broj kompliciranih apendicitisa imao stabilnu incidenciju (71).

Imajući u vidu klinički značaj, akutni apendicitis se dijeli na nekomplicirani koji se definira kao upalno promijenjeni crvuljak bez znakova nekroze ili perforacije, i na komplicirani u kojem se pronalaze znakovi fokalne ili transmuralne nekroze s ili bez perforacije, intraabdominalni apsces, periapendikularna zadržana flegmona ili postojanje gnojne slobodne tekućine (72). Patohistološki, u flegmonoznom (supurativnom) apendicitisu može se pronaći neutrofilna infiltracija sluznice, podsluznice i mišićnog sloja, uz transmuralnu upalu, opsežnu ulceraciju i intramuralne apscese s vaskularnom trombozom. Ovaj oblik apendicitisa možda neće biti očit makroskopski, osim u dobro razvijenim slučajevima kada se mogu primijetiti zamućena seroza, dilatacija i kongestija površinskih krvnih žila ili fibrinopurulentni serozni eksudat. Osnovna značajka gangrenoznog, uključujući i perforirani, apendicitisa je nekroza stijenke. U ovom slučaju pronalazi se transmuralna upala s područjima nekroze i opsežne ulceracije sluznice. Makroskopski stijenka može biti ljubičaste, zelene ili crne boje (73).

1.1.5.3. Dijagnoza apendicitisa

Klinička dijagnoza akutnog apendicitisa je često izazovna i zahtjeva kombinaciju kliničkih, laboratorijskih i radioloških nalaza te bi se mogla unaprijediti korištenjem kliničkih sustava bodovanja (62).

1.1.5.3.1. Klinički pokazatelji

Klinički znakovi i simptomi bolesnika su varijabilni i inicijalno suptilni. Klasična klinička prezentacija započinje slabo definiranom boli u trbuhu, lokaliziranom epigastrično, periumbilikalno ili difuzno. Početak boli obično je brzo praćen mučninom. U otprilike 50-60% bolesnika s apendicitisom, koji se prezentiraju s periumbilikalnom bolnošću, dolazi do migracije boli u donji desni abdominalni kvadrant unutar 24 sata. Također dolazi i do inapetencije u 80-85% bolesnika, a 40-60% njih prijavljuje mučninu s ili bez povraćanja. Međutim, apendicitis je malo vjerojatan kod osoba s normalnim apetitom. U slučaju bolova u trbuhu i povraćanja, bitno je ispitati redoslijed događaja jer pojava boli prije povraćanja povećava vjerojatnost

akutnog apendicitisa. Ako mučnina i povraćanje prethode boli, iskusni kliničar trebao bi razmotriti drugu dijagnozu, poput gastroenteritisa. Čest je i subfebrilitet, dok više temperature i treskavice mogu sugerirati druge dijagnoze, kao i perforaciju apendicitisa ili druge neapendikularne izvore boli. U slučaju povišene tjelesne temperature iznad 38,3 °C trebalo bi posumnjati na perforaciju (67, 74-76). Inicijalnu bol bolesnici osjećaju zbog visceralne inervacije, a kasnija lokalizirana bol nastaje zbog nadražaja parijetalnog peritoneuma uslijed progresije upalnog procesa. Konstipacija i mučnina praćena profuznim povraćanjem mogu ukazivati na razvoj generaliziranog peritonitisa nakon perforacije, ali rijetko imaju bitnu ulogu u jednostavnim upalama (76).

Ovisno o lokaciji crvuljka, bolesnici prijavljuju različite simptome. Bolesnici se mogu žaliti na bolove desno lumbalno ili u kostovertebralnom kutu u slučaju retrocekalnog crvuljka, a u muškaraca i bol u desnom testisu. Zdjelični ili retroilealni položaj odražavat će se zdjeličnom, rektalnom ili adneksalnom boli, a rijetko i boli u lijevom donjem kvadrantu. Dok kod subcekalnog ili zdjeličnog mogu prevladavati suprapubična bol i dizurične tegobe (76).

Prilikom inspekcije bolesnika zamjećuje se dojam blaže bolesti. Često leže mirno kako bi izbjegli iritaciju parijetalnog peritoneuma koja se pojavljuje kretanjem (67). U kombinaciji s anamnezom, klinički pregled može biti dovoljan za postavljanje dijagnoze, pogotovo kod muškaraca kod kojih diferencijalna dijagnoza nije toliko opsežna kao kod žena (22).

Trbuh je najčešće mekane stijenke s bolnom osjetljivošću oko McBurneyjeve točke i lokaliziranom napetošću stijenke. Razlika između aksilarne i rektalne temperature više od 1 °C sugerira zdjeličnu upalu koja može biti ili zbog apendicitisa ili drugih zdjeličnih upalnih stanja. Nevoljno napinjanje mišića u donjem desnom kvadrantu česta je pojava i obično prethodi peritonealnom nadražaju (76). Palpabilna masa rijetko je prisutna, a kada jest, može biti teško uočjiva zbog nelagode prilikom duboke palpacije i debljine trbušne stijenke. Prisutnost ovog nalaza može sugerirati apscesnu kolekciju ili peritiflitičku flegmonu, što odražava perforaciju crvuljka s adherencijom omentuma i okolnog tkiva (22). Među bolesnicima s retrocekalno smještenim crvuljkom, bol ne mora biti lokalizirana sve dok se ne razvije peritonitis zbog perforacije. Također, polakisurija ili učestalije stolice mogu biti prisutne zbog upale koja nadražuje susjedni mjehur ili debelo crijevo (67). Bitniji simptomi i znakovi s osjetljivošću i specifičnošću za apendicitis su prikazani u **Tablici 3.** (74).

Digitorektalni pregled ne doprinosi značajno dijagnostičkoj točnosti, stoga bi se trebao raditi samo kod onih za koje se sumnja na zdjeličnu ili materničnu patologiju, ili u atipičnim kliničkim slikama koje sugeriraju na zdjelični ili retrocekalni apendicitis (76). Ginekološki pregled s određivanjem β -hCG-a je bitan kod svih žena reproduktivne dobi sa sumnjom na

apendicitis zbog isključenja moguće trudnoće kao i ginekoloških upalnih zbivanja te moguće daljnje dijagnostičke obrade poput CT-a abdomena. (72, 74). Osim navedenog, rijetke atipične kliničke slike apendicitisa uključuju i bol u gornjem lijevom kvadrantu koja se pojavljuje kod bolesnika s malrotacijom crijeva ili *situsom inversusom* (66). Apendicitis se može pojaviti u sklopu nekoliko specifičnih entiteta, uključujući Amyandovu ingvinalnu kilu, de Garengeotovu femoralnu kilu te Spigelijanovu kilu (24).

Tablica 3. Bitniji simptomi i znakovi u akutnom apendicitisu

Simptomi i znakovi	Osjetljivost	Specifičnost	Napomena
Temperatura	60-75%	65-75%	nespecifičan simptom
Bol u donjem desnom kvadrantu	80-90%	40-60%	najbitniji simptom koji je prisutan u više od 90% bolesnika
Mučnina	58-68%	40-50%	Učestalost od 80-90%
Povraćanje	50-60%	45-69%	Učestalost 75%
Pojava boli prije povraćanja	90-100%	55-70%	Učestalost veća od 90%
Inapetencija	80-90%	55-70%	Učestalost 90%
Migracija boli	55-75%	70-90%	Učestalost od 50%
McBurneyev znak	50-94%	75-86%	Najbitniji znak
Rovsingov znak	22-68%	58-96%	Odsustvo ovoga ne isključuje dijagnozu apendicitisa
Znak psoasa	13-42%	79-97%	Povezan s položajem crvuljka retroperitonealno i retrocekalno
Znak opturatora	5-15%	90-95%	Povezanost s položajem crvuljka u zdjelici
Nevoljno napinjanje mišića	39-74%	57-84%	povezanost s lokaliziranom upalom peritoneuma
Povratna osjetljivost	55-70%	65-80%	povezanost s upalom peritoneuma
Rigiditet	20-30%	80-90%	povezanost s intraabdominalnom infekcijom

Trudnice su dijagnostički izazov kod akutnog apendicitisa. Anatomske i fiziološke promjene povezane s trudnoćom moraju se uzeti u obzir prilikom uzimanja anamneze i fizikalnog pregleda. Tijekom prvih 6 mjeseci trudnoće, simptomi i znakovi akutnog apendicitisa su isti kao kod žena koje nisu trudne, no dijagnostičke poteškoće ostaju iz više razloga. Postoji prigušenje simptoma i znakova zbog distenzije trbuha, drukčijeg položaja intraabdominalnih organa te smanjenog odgovora tkiva na upalu. U slučaju da je cekum mobilan može doći do promjene položaja i crvuljka skupa s njim. Također, mučnina, povraćanje i bolovi u trbuhu koji su normalni u trudnoći, a osobito u prvom tromjesječju, dovode do dijagnostičkih poteškoća. Atipična klinička slika najčešće je prisutna u drugom

tromjesečju. Pa tako bol u gornjem desnom kvadrantu, kontrakcije maternice, dizurija i proljev mogu prikrivati akutni apendicitis (77). U gerijatrijskoj populaciji najčešći simptomi apendicitisa su bol u donjem dijelu trbuha (90%) i inapetencija (60%). Nadalje, kod pacijenata koji nisu u mogućnosti komunicirati, klinička slika može biti još suptilnija. Rana dijagnoza na temelju samo na temelju znakova i simptoma otežana je u ovoj populaciji, što rezultira većom stopom perforacije i mortaliteta, stoga je potrebna pažljiva evaluacija (78).

Također, potrebno je imati na umu da prethodna apendektomija ne isključuje definitivno dijagnozu apendicitisa, jer, iako rijetko, opisano je stanje poznato kao „*stump*“ apendicitis (67). To je rijetka komplikacija apendektomije, prvi put opisana 1945. godine, u kojoj dolazi do upale ostatnog dijela crvuljka uslijed nepotpune resekcije crvuljka. Ovo stanje nije često uključeno kao diferencijalna dijagnoza kod osoba s bolovima u trbuhu, a koje su prethodno apendektomirane. Klinička slika može odgovarati onoj u akutnom apendicitisu, a najčešći simptom je bol lokalizirana u donjem desnom kvadrantu trbuha. Klinički značaj leži u tome što se teže dijagnosticira, što dovodi do odgode u liječenju te posljedično porastu morbiditeta (79, 80).

1.1.5.3.2 Biokemijski markeri

Postoji mnogo biomarkera koji se trenutno koriste u svakodnevnoj kliničkoj praksi, evaluiraju se kroz klinička ispitivanja ili se još istražuju na eksperimentalnoj razini. Neki od njih uključuju broj leukocita, C-reaktivni protein (CRP), broj neutrofila kao postotak od leukocita ili u apsolutnoj vrijednosti, omjer neutrofila i limfocita (engl. *neutrophil-to-lymphocyte ratio*, NLR), interleukini, bilirubin, prokalcitonin i kalprotektin (81). Vrijednosti leukocita, neutrofila i CRP-a imaju suboptimalne osjetljivosti i specifičnosti pa se ne mogu same koristiti u dijagnostici apendicitisa. Vrijednosti leukocita i CRP-a koriste se i u kliničkim sustavima bodovanja koji se koriste u stratifikaciji rizika za dijagnozu apendicitisa. Stoga se radi na pokušajima identifikacije novih biomarkera koji bi bili specifični za apendicitis, a koristili bi se ili samostalno ili u kombinaciji s drugim biomarkerima (82). Vrijednosti serumskog bilirubina su razmatrane kao mogući marker perforiranog apendicitisa. Do povišenih vrijednosti bilirubina dolazi zbog bakterijemije što dovodi do endotoksemije i posljedično smanjene sekrecije bilirubina iz žučnih kanala. Međutim, sama hiperbilirubinemija ima nisku preciznost u dijagnosticiranju kompliciranog apendicitisa s očekivanom perforacijom. U slučajevima gdje su vrijednosti bilirubina povišene, vjerojatnija je dijagnoza kompliciranog apendicitisa, dok normalne vrijednosti bilirubina idu u prilog nekompliciranom apendicitisu (83). Također, široko korišteni upalni biomarker CRP nije specifičan za

apendicitis. Njegove vrijednosti porastu 8-12 sati nakon početka upalnog procesa, pa se može koristiti kao prediktor uznapredovale upale radije nego u ranoj dijagnostici jednostavnog apendicitisa (82).

1.1.5.3.3 Radiološke metode

Prema rezultatima meta-analize, CT ima najveću osjetljivost (>80%) i specifičnost (>93%) kod odraslih bolesnika u usporedbi s drugim radiološkim pretragama, dok je u pedijatrijskoj populaciji prisutna velika varijabilnost u rezultatima (84). U WSES smjernicama preporučuje se rutinski napraviti UZV u kombinaciji s kliničkim parametrima, nakon inicijalne procjene i stratifikaciji rizika koristeći kliničke sustave bodovanja, kako bi se poboljšala osjetljivost i specifičnost te smanjila potreba za CT-om. Također, kod osoba mlađih od 40 godina s visokim rizikom od apendicitisa prema kliničkim sustavima bodovanja, može se izbjeći MSCT prije dijagnostičko-terapijske laparoskopije. Ultrazvuk (*point of care ultrasound, POCUS*) se preporučuje kao prvi radiološki dijagnostički alat i kod djece i odraslih. U slučaju negativnog UZV nalaza, preporučuje se se niskodozni (engl. *low-dose*) CT s kontrastom umjesto standardnog CT-a s kontrastom kod odraslih, dok se kod djece kao idući korak preporuča daljnja radiološka obrada koja će ovisiti o lokalnoj dostupnosti i stručnosti osoblja. Ako se simptomi ne smiruju kod bolesnika kod kojeg nije vjerojatna dijagnoza apendicitisa, preporučavaju se presječne slikovne metode (engl. *cross sectional imaging*) prije operacije, a na koncu i eksplorativna laparoskopija u slučaju perzistiranja tegoba i negativnog radiološkog nalaza. Kod trudnica prvi je izbor transabdominalni ultrazvuk tehnikom dozirane kompresije (engl. „*graded compression*“) kod sumnje na apendicitis, a u slučaju negativnog UZV-a preporuča se napraviti MR ako je dostupan (62).

1.1.5.3.4 Klinički sustavi bodovanja

Različiti klinički sustavi bodovanja razvijeni su kako bi pomogli u dijagnozi akutnog apendicitisa, od kojih je najpoznatiji Alvarado bodovni sustav (85). Prema WSES smjernicama iz 2020. godine preporučuje se koristiti AIR („*Appendicitis Inflammatory Response*“) bodovni sustav i PAS („*pediatric appendicitis score*“) bodovni sustav kao kliničke prediktore akutnog apendicitisa (62). Oni se koriste anamnestičkim podacima, kliničkim pokazateljima i laboratorijskim nalazima kako bi se procijenio rizik apendicitisa. Njihova usporedba prikaza je u **Tablici 4**. Razvijen je i RIPASA bodovni sustav zbog bolje diferencijacije bolesti u azijskoj populaciji, koji nadmašuje Alvarado u sposobnosti detekcije bolesti (85).

Osim ovih, razvijeni su i brojni drugi sustavi bodovanja u predikciji akutnog apendicitisa (86). Također, postoje razni drugi modeli vezani za akutni apendicitis, primjerice u svrhu predikcije proširene resekcije u kirurškom liječenju apendicitisa ili neuspjeha konzervativnog liječenja (87, 88).

Tablica 4. Usporedba Alvarado, PAS i AIR sustava bodovanja

	Alvarado		PAS		AIR	
Leukociti	> 10 000	2	>10 000	1	10-14 000	1
					≥15 000	2
Neutrofili	>75%	1	>75%	1	70-84%	1
					≥85%	2
CRP					10-49g/l	1
					≥50g/l	2
Temperatura	>37,3C	1	>38C	1	>38,5C	1
			Prilikom		Slaba	1
Povratna osjetljivost	Da	1	kašlja,	2	Srednja	2
			perkusijske,		Jaka	3
			skakanja			
Inapetencija	Da	1	Da	1		
Mučnina/povraćanje	Da	1	Da	1	Samo povraćanje	1
Migracija	Da	1	Da	1		
Bol u DDK					Da	1
Ukupni zbroj		10		10		12
Niska vjerojatnost		1-4		1-3		0-4
Srednja vjerojatnost		5-6		4-7		5-8
Visoka vjerojatnost		7-10		8-10		9-12

DDK – donji desni abdominalni kvadrant

AIR sustav bodovanja je razvijen s ciljem stvaranja kliničkog sustava bodovanja koji je lako primjenjiv u praksi, a temeljen je na prospektivno prikupljenim podacima pacijenata koji su obrađivani zbog suspektog apendicitisa. Kliničke varijable su gradirane prema težini simptoma i znakova umjesto binarne podjele, a laboratorijske vrijednosti su podijeljene u intervale (89). Definirane su tri dijagnostičke zone: jedna s visokom osjetljivošću za apendicitis, namijenjena identifikaciji pacijenata koji mogu biti sigurno otpušteni uz ambulantno praćenje; druga s visokom specifičnošću za prepoznavanje pacijenata koji mogu biti operirani bez dodatnih pretraga; te treća skupina nejasnih slučajeva kojima je potrebna dodatna dijagnostička obrada odnosno skupine niske, srednje i visoke vjerojatnosti (89). Na provedenoj validaciji autori su objavili da AIR sustav bodovanja ima dobre performanse u

detekciji kompliciranog apendicitisa u odnosu na općenito apendicitis, kao i dobre performanse u detekciji apendicitisa kod mlađih od 15 godina i kod pacijenata s trajanjem simptoma dužim od 47 sati (90).

1.1.5.4. Diferencijalna dijagnoza

Akutni apendicitis predstavlja dijagnostički izazov zbog nespecifičnosti simptoma i znakova, što rezultira mnogim potencijalnim diferencijalno-dijagnostičkim uzrocima boli koji mogu biti abdominalni, kao i neabdominalni. Iako su određene dijagnoze vjerojatnije u određenim kliničkim slučajevima. Primjerice, kod Meckelovog divertikulitisa, koji može uzrokovati slične simptome, može se anamnestički saznati o epizodama bezbolne hematokezije. A kod gastroenteritisa, koji je izrazito češći uzrok, dolazi do mučnine i povraćanja koji prethode bolovima u trbuhu, ili pak profuznoj dijareji. Crohnova bolest koja zahvaća terminalni ileum može u početnoj prezentaciji nalikovati apendicitisu, no detaljnijom anamnezom često se otkriva subakutni tijek s vrućicom, gubitkom težine i bolovima. Kod pacijenata srednje i starije životne dobi, diferencijalno-dijagnostički treba razmotriti i druga upalna stanja poput želučanog ili duodenalnog ulkusa, kolecistitisa i pankreatitisa. Nadalje, simptomi cekalnog ili sigmoidnog divertikulitisa mogu se preklapati s onima akutnog apendicitisa. Bolesnici s mobilnim sigmoidnim kolonom mogu imati bolove u donjem desnom kvadrantu zbog mobilnog, dekstroponiranog sigmoidnog debelog crijeva. Međutim, ovi bolesnici navode bržu progresiju do lokaliziranog peritonealnog nadražaja te promjene u pražnjenju crijeva. Maligne bolesti također mogu uzrokovati akutnu bol u desnom donjem kvadrantu zbog perforacije karcinoma cekuma ili apendicitisa uzrokovanog tumorskom opstrukcijom apendikularnog ušća. U žena reproduktivne dobi, dijagnosticiranje podležećeg uzroka boli u desnom donjem kvadrantu može biti još izazovnije. Osim prethodno navedenih uzroka, mlade žene mogu imati bol uzrokovanu opstetričkim i ginekološkim etiologijama poput rupture ovarijske ciste ili folikula, torzije jajnika, ektopične trudnoće, akutnog salpingitisa i tuboovarijalnog apscesa. Detaljna anamneza, uključujući nedavnu menstrualnu povijest, kao i ginekološki pregled, mogu biti od pomoći u razlikovanju ovih uzroka boli od akutnog apendicitisa. Unatoč tome, dijagnoza apendicitisa može biti otežana u ovoj populaciji pacijenata, s višim stopama pogrešne dijagnoze koje su opisane kod žena reproduktivne dobi (67).

1.1.6. Liječenje

Nakon što je postavljena dijagnoza apendicitisa, potrebno je odlučiti o daljnjem tijeku liječenja. Kontroverza o najadekvatnijem načinu liječenja i dalje perzistira, pa se, ovisno radi li se o nekomplikiranom ili kompliciranom apendicitisu, može pristupiti konzervativnom ili operacijskom liječenju. Osim toga, po završetku konzervativnog liječenja može se pristupiti i intervalnoj apendektomiji, ali načini liječenja su varijabilni i ne postoji univerzalni konsenzus (84). Također, različiti modaliteti liječenja primjenjuju se ovisno radi li se o djeci, odraslima, osobama starije životne dobi, imunokompromitiranim ili trudnicama (91).

1.1.6.1. Konzervativni pristup

Još od 1950-ih liječnici su pokušavali konzervativnim metodama liječiti apendicitis. Uzimajući u obzir opisane razlike između nekomplikiranog i kompliciranog apendicitisa, kao i na imunološku funkciju crvuljka te na njegovu ulogu u crijevnoj flori, postoji povećan značaj očuvanja crvuljka (92). Osim toga, postoje stanja zbog kojih bi očuvanje crvuljka moglo biti korisno; primjerice, apendikularni konduiti bi mogli biti korišteni u liječenju uroloških, neuroloških i bilijarnih stanja. Međutim, razvojem kirurških i minimalno invazivnih tehnika ovaj način korištenja crvuljka danas se rijetko koristi (93).

Konzervativno liječenje podrazumijeva primjenu odgovarajućih antibiotika i hiperhidraciju bolesnika kako bi se izbjeglo operacijsko liječenje. Sami antibiotici se mogu primjenjivati peroralno ili intravenski. Kod ovog načina liječenja pristupa se apendektomiji samo kod osoba s lošim kliničkim odgovorom na konzervativne mjere liječenja, nastavkom pogoršanja stanja unatoč antibiotskoj terapiji, ili kod onih s ponavljajućim apendicitisom (91). Prema WSES smjernicama može se pokušati antibiotska terapija kao sigurna alternativa kod određenih bolesnika s nekomplikiranim apendicitisom i odsustvom apendikolita, kako kod djece tako i kod odraslih. Ovaj pristup se ne preporučuje tijekom trudnoće (62).

Odabir antibiotske terapije temelji se na empirijskom pristupu, uzimajući u obzir očekivane mikroorganizme uključene u patogenezu akutnog apendicitisa, kao i težinu intraabdominalne infekcije. Dodatno, u obzir se uzima prisutnost rizičnih čimbenika za razvoj antibiotske rezistencije, i/ili prethodna terapijska neučinkovitost antibiotika za prijašnje infekcije, kao i neke specifične okolnosti poput nedavnih putovanja u područja s visokom prevalencijom rezistentnih patogena, poznate kolonizacije takvim organizmima, imunosupresije ili značajnih komorbiditeta (74). Nekomplirani i komplicirani apendicitis, u kontekstu intraabdominalnih infekcija stečenih u zajednici, smatraju se infekcijama blagog do

umjerenog intenziteta kod pacijenata bez predispozicije za antibiotsku rezistenciju ili terapijski neuspjeh. Stoga se u većini slučajeva primjenjuju antibiotici učinkoviti protiv streptokoka i osjetljivih *Enterobacteriaceae* te anaerobnih uzročnika. Kod visokorizičnih bolesnika indicirana je empirijska antibiotska terapija širokog spektra koja obuhvaća *Pseudomonas aeruginosa* i *Enterobacteriaceae* (74).

Meta analiza iz 2019. godine, koja je uspoređivala antibiotsko liječenje i apendektomiju za nekomplikirani akutni apendicitis u odraslih i djece, pronašla je veću stopu učinkovitosti liječenja te veću stopu uspjeha liječenja bez razvoja komplikacija u jednogodišnjem praćenju. Do neuspjeha antibiotskog liječenja došlo je u 8,5% slučajeva, a stopa povrata bolesti iznosila je 19,2%. Po ovim rezultatima, autori zaključuju kako je antibiotsko liječenje izvediv način liječenja akutnog nekomplikiranog apendicitisa (94). Druga meta-analiza 2024. godine, koja se fokusirala na randomizirana kontrolna istraživanja, potvrdila je kako kirurško liječenje ima bolju uspješnost u jednogodišnjem praćenju, kao i uspješnost liječenja bez razvoja komplikacija. Naspram tome, nije pronađena statistički značajna razlika u duljini trajanja hospitalizacije, dok su troškovi liječenja išli u prilog antibiotskom liječenju (95). Isto tako, u recentnom Cochraneovom sustavnom pregledu uspoređivani su rezultati randomiziranih kontrolnih istraživanja u kojima su većina ispitanika bile odrasle osobe s nekomplikiranim apendicitisom. Pronađeno je da u skupini koja je liječena antibioticima u 76 na 1000 osoba češće dolazi do neuspjeha liječenja iako razlika ne mora nužno biti klinički značajna. Nakon provedenog antibiotskog liječenja 30,7% osoba na kraju je trebalo kirurško liječenje, ali dokazi su nepouzdana (96). Međutim, neka manja randomizirana kontrolna istraživanja dala su oprečne rezultate o korisnosti antibiotske terapije naspram placebo u konzervativnom liječenju nekomplikiranog apendicitisa. Dva nisu potvrdila bolji uspjeh antibiotskog liječenja u odnosu na ne-antibiotski pristup, dok je jedno pokazalo da je antibiotsko liječenje uspješnije u prevenciji kirurške eksploracije i apendektomije. Ovdje je važno napomenuti razliku u metodologiji. U prva dva istraživanja uključeni su ispitanici s CT potvrđenom dijagnozom apendicitisa, dok u posljednjem nisu svi ispitanici imali nalaz MSCT-a (97-99).

1.1.6.2. Kirurški pristup

Laparoskopska apendektomija je sigurna tehnika zahvata kod većine bolesnika. Međutim, otvoreni zahvat može biti bolji izbor kod bolesnika s opsežnim priraslicama od prethodnih zahvata ili kod onih koji ne mogu podnijeti pneumoperitoneum zbog drugih komorbiditeta. Isto tako, bolesnici s ozbiljnim komorbiditetima ili oni na antikoagulantnoj

terapiji mogu biti kandidati za inicijalno intravensko antibiotsko liječenje. U uznapredovanoj fazi bolesti i prisutnom difuznom intraabdominalnom kontaminacijom, bolesnici bi mogli imati korist od klasičnog operacijskog liječenja zbog bolje lavaže trbušne šupljine. Kod pacijenata koji se prezentiraju sa septičkim šokom uslijed perforacije i apscesa, nužna je brza kontrola izvora infekcije. Ukoliko je dostupna, hitna drenaža vođena slikovnim metodama pod lokalnom anestezijom može biti korisna. Međutim, sepsa se ne bi trebala smatrati kontraindikacijom za hitno operacijsko liječenje, bilo laparoskopski ili klasično, pod uvjetom da se tim pacijentima pruži odgovarajuća intravenska nadoknada tekućine i intravenski antibiotici preoperativno (100). Komplikacije laparoskopske apendektomije možemo podijeliti na intraoperativne, kao što su krvarenje (najčešće iz apendikularne arterije ili mjesta insercije troakara), ozljeda crijeva, curenje gnoja ili fekolita iz crvuljka, nekompletna apendektomija; te postoperativne kao što su intraabdominalni apscesi, curenje sadržaja iz bataljka crvuljka, upala kirurške rane u vidu celulitisa ili apscesne kolekcije, opstrukcije tankog crijeva i fistule (100, 101).

Otvorena apendektomija se uobičajeno izvodi pod općom anestezijom, premda je moguća primjena i regionalne anestezije. Incizija se najčešće izvodi na McBurneyevoj točki, kosim rezom (McBurneyeva incizija) ili poprečnim rezom (Rocky-Davisova incizija). U slučajevima perforiranog apendicitisa s peritiflitičkom flegmonom, prikladna je donja medijana laparotomija (65).

Brojni su radovi objavljeni na temu usporedbe klasične i laparoskopske apendektomije. Zaključci ovih analiza ukazuju na sljedeće: kod klasičnog pristupa trajanje operacijskog zahvata je kraće i manja je stopa postoperativnih intraabdominalnih apscesa, dok laparoskopski pristup rezultira manjom postoperativnom boli, kraćim boravkom u bolnici, bržim povratkom na posao i manjom učestalošću infekcija rane. Laparoskopska apendektomija može biti posebno korisna za određene skupine pacijenata, uključujući žene reproduktivne dobi, pretila pacijente i starije osobe. Kod žena reproduktivne dobi, laparoskopija omogućuje istovremenu dijagnostiku i terapiju, izbjegavajući nepotrebnu laparotomiju u slučaju otkrivanja neginekološke patologije. Kod pretilih pacijenata, laparoskopski pristup pokazao je značajno smanjenje morbiditeta u usporedbi s otvorenim pristupom. Kod starijih pacijenata, laparoskopska apendektomija povezana je s nižom smrtnošću kod nekomplikiranog apendicitisa i manje kompliciranim postoperativnim tijekom kod perforiranog apendicitisa. Studije provedene na pedijatrijskoj populaciji ukazuju na nižu incidenciju infekcija kirurške rane, manju potrebu za analgeticima i brži oporavak kod laparoskopske apendektomije u usporedbi s otvorenom. Vrijeme operacije bilo je duže kod laparoskopskog pristupa samo za komplicirane slučajeve apendicitisa. Zaključno, odluka o kirurškom pristupu apendektomiji

treba se temeljiti na individualnim čimbenicima pacijenta i iskustvu kirurga s određenom tehnikom (67).

Postoje i neke novije tehnike, kao što je SILS, u kojoj se koristi samo jedan radni troakar. Međutim, nisu pokazani bolji kozmetički učinci; osim toga, moguća je i veća stopa incizijskih hernija postoperativno. Endoskopska kirurgija kroz prirodne tjelesne otvore (engl. *natural orifice tranluminal endoscopic surgery*, NOTES) se pokazala superiornom u smislu boljih kozmetičkih učinaka i postoperativne boli; međutim, potrebna su dodatna istraživanja kako bi potvrdila ove rezultate. Robotskim pristupom operater ima puno bolji opseg kretnji instrumentima, ali je ova metoda izrazito skupa i često zahtijeva veće troakare; stoga se ne koristi u svakodnevnoj kliničkoj praksi.

Apendektomije također imaju kliničku primjenu u obliku intervalne, incidentalne i profilaktičke apendektomije. U suštini, sva tri oblika zahvata provode se u preventivne svrhe (93).

Apendektomija po smirenju upale nakon početne kliničke slike akutnog apendicitisa se naziva odgođena ili intervalna apendektomija. Često se izvodi u slučajevima nakon zadržane perforacije crvuljka ako je bilo odlučeno pristupiti inicijalno konzervativnom liječenju. Također, intervalnom apendektomijom se može isključiti neoplazma kao uzrok inicijalne upale (91). U slučaju intraoperativnog nalaza morfološki urednog, inocentnog, crvuljka kod bolesnika kojemu je indiciran operacijski zahvat zbog sumnje na apendicitis potrebno je svakako učiniti apendektomiju ako se ne pronađe drugo patološko stanje, uzimajući u obzir visoku varijabilnost i netočnost procjene operatera o ranoj fazi upale (62). Svakako, potrebno je napraviti i inspekciju desnog kolona zbog mogućih divertikula, inspekciju tankog crijeva zbog mogućeg Meckelovog divertikula, a kod žena i jajovoda te jajnika (100).

Incidentalna apendektomija definira se kao izvođenje apendektomije u sklopu drugog kirurškog zahvata. Razlozi zbog kojih se ovaj pristup ponekad provodi su sprečavanje mogućih komplikacija apendicitisa, moguće poteškoće u diferencijalnoj dijagnozi i izbjegavanje potencijalne druge operacije u budućnosti. Prema nekim epidemiološkim istraživanjima, zabilježeno je da 20-25 incidentalnih apendektomija sprječava jedan budući slučaj apendicitisa i moguće komplikacije (93).

Nadalje, profilaktička apendektomija kao elektivni zahvat u određenih bolesnika može se izvesti kao dijagnostički ili terapijski zahvat zbog verificiranih tvorbi crvuljka, stranih tijela ili drugih stanja (93). Zanimljivo je da u odrađenim situacijama profilaktička apendektomija je obavezna, primjerice prilikom ekspedicija na Antarktiku, ovisno o propisima zemlje odakle ljudstvo dolazi (102). Međutim, istraživanje koje je koristilo probabilističke modele za

potencijalnu 2,5-godišnju misiju na Mars, nije pronašlo prednosti profilaktičke apendektomije za svemirske misije uzimajući u obzir potencijale rizike kao što su opstrukcije crijeva zbog priraslica (103).

1.2. Strojno učenje

Strojno učenje (engl. *machine learning*, ML) je grana umjetne inteligencije (engl. *artificial intelligence*, AI) koja se bavi primjenom algoritama koji mogu definirati svoja pravila na temelju unesenih podataka kroz iterativno treniranje i unapređivanje bez eksplicitnog programiranja (104). To znači da je potrebna intervencija korisnika u fazi unosa podataka, odabira značajki i za verifikaciju izlaznih varijabli, dok algoritmi sami upravljaju oblikovanjem značajki i procesom učenja (105).

U osnovi, strojno učenje obuhvaća nadzirano (engl. *supervised learning*), nenadzirano (engl. *unsupervised learning*), polunadzirano (engl. *semi-supervised learning*) i podržano učenje (engl. *reinforcement learning*) (Slika 2.) (106).

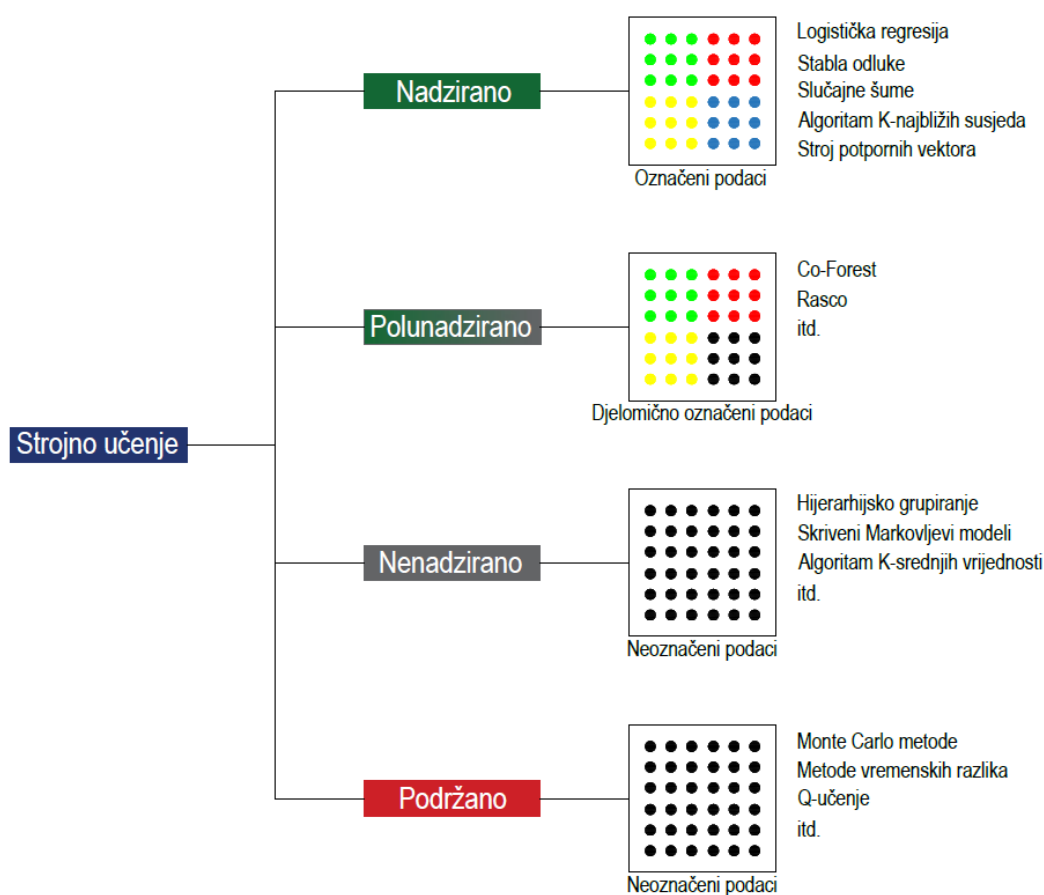
U nadziranom učenju algoritam koristi označene podatke za obuku modela. Za vrijeme treniranja algoritmi se uče prepoznavanju obrazaca i pravilnosti na temelju označenih ulaznih i izlaznih varijabli kako bi mogli izraditi model primjenjiv na buduće ulazne podatke. Ovi ML modeli se mogu dalje podijeliti ovisno o tome kakva je varijabla koju predviđamo. U klasifikacijskim modelima predikcije se rade za kategorijske podatke, a u regresijskim modelima predikcije se rade za numeričke varijable (107). Trening i odabir modela trebaju se provesti na podacima koji neće biti korišteni prilikom testiranja konačne izvedbe modela kako ne bi došlo do „curenja“ podataka i posljedično prenaučenosti modela (engl. *overfitting*). Uobičajeni pristup podrazumijeva podjelu podataka na dva ili više skupova. Tradicionalno koristimo skup za trening modela („trening“ koji služi za učenje algoritma), validacijski set (za optimizaciju hiperparametara i validaciju rezultata) i testni set (za završnu procjenu učinkovitosti modela) (108, 109).

Nenadzirano strojno učenje ne oslanja se na oznake podataka, već isključivo na vrijednosti podataka, tj. na intrinzičnu strukturu podataka. Za razliku od nadziranih modela, ovdje se ne pokušava predvidjeti ishod, nego se podaci grupiraju u klastere po sličnosti (110). Stoga je njihova glavna zadaća grupirati podatke u odgovarajuću grupu. Ovi modeli se mogu koristiti za otkrivanje neprepoznatih uzoraka ili unutarnjih struktura (107).

Podržano učenje je računalni pristup za razumijevanje i automatiziranje učenja temeljenog na ciljevima i donošenju odluka. Model ne zna koje korake treba napraviti, već

mora otkriti koji koraci donose najviše nagrade. Razlikuje se od ostalih računalnih pristupa po tome što se usredotočuje na učenje kroz izravnu interakciju s okolinom, bez potrebe za detaljnim nadziranjem ili potpunim razumijevanjem okoline (111).

Polunadzirano strojno učenje ima mogućnost učenja iz male količine dostupnih podataka za trening i istovremeno iz veće količine neoznačenih podataka. Kombiniranjem ovih pristupa mogu se dobiti modeli visoke preciznosti uz znatno manju količinu podataka za trening. Osim navedenog, prednost polunadziranog strojnog učenja je izbjegavanje potrebe za kodiranje cijelog skupa podataka (112).

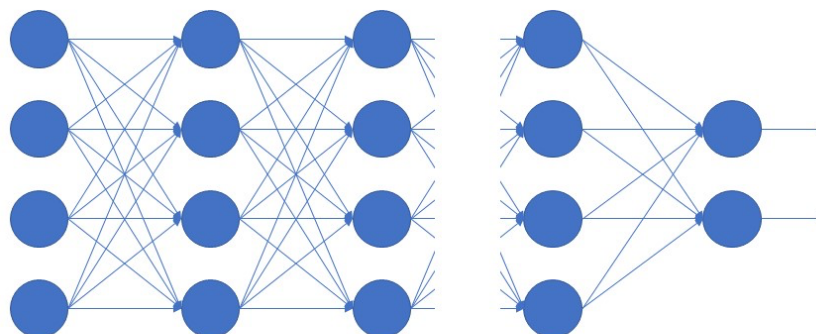


Slika 2. Prikaz podjele algoritama strojnog učenja. (Preuzeto i prilagođeno prema: Rafique R, Islam SMR, Kazi JU. Machine learning in the prediction of cancer therapy. Comput Struct Biotechnol J. 2021;19:4003-4017. Djelo otvorenog pristupa licencirano pod CC-BY 4.0 licencom)

Duboko učenje (engl. *deep learning*, DL) je podskup strojnog učenja koje koristi međusobno povezane jedinice koje se nazivaju umjetne neuronske mreže (engl. *artificial neural networks*, ANN). Organizacija ANN-a je modelirana prema biološkim neuronskim strukturama - neuronima. U DL-u jednostavne računalne jedinice nazivaju se neuronima; one

dodjeljuju različite vrijednosti unesenim podacima te ih potom proslijeđuju kroz matematički algoritam nazvan aktivacijska funkcija (engl. *activation function*). Filtracijom podataka kroz ove mreže dolazi se do izlaznih vrijednosti. ANN-ovi organiziraju neurone u višeslojnu arhitekturu koja uzima zadanu vrijednost i obrađuje je kroz ulazni sloj (engl. *input layer*), zatim kroz više skrivenih slojeva gdje se podaci procesiraju bez da su direktno vidljivi (engl. *hidden layer*), te potom prolaze kroz završni sloj (engl. *output layer*) koji daje konačni rezultat. Stoga, DL opisuje klasu ANN-ova koji postaju „duboki“ kada se sastoje od mnogih skrivenih slojeva (113).

Za većinu zadataka, ANN-ovi proslijeđuju informacije u jednom smjeru, što je poznato kao *feedforward* neuronska mreža (**Slika 3.**). To bi značilo da je podatak iz svakog čvora u sloju dobio vrijednost iz čvora u prethodnom sloju, obradio je, i potom proslijedio u čvor u idućem sloju. Osim toga, postoje i rekurentne neuronske mreže, u kojima se podaci mogu prenositi između čvorova unutar sloja ili prema prethodnom sloju kada se vrijednosti ponovno obrađuju i proslijeđuju dalje (114).



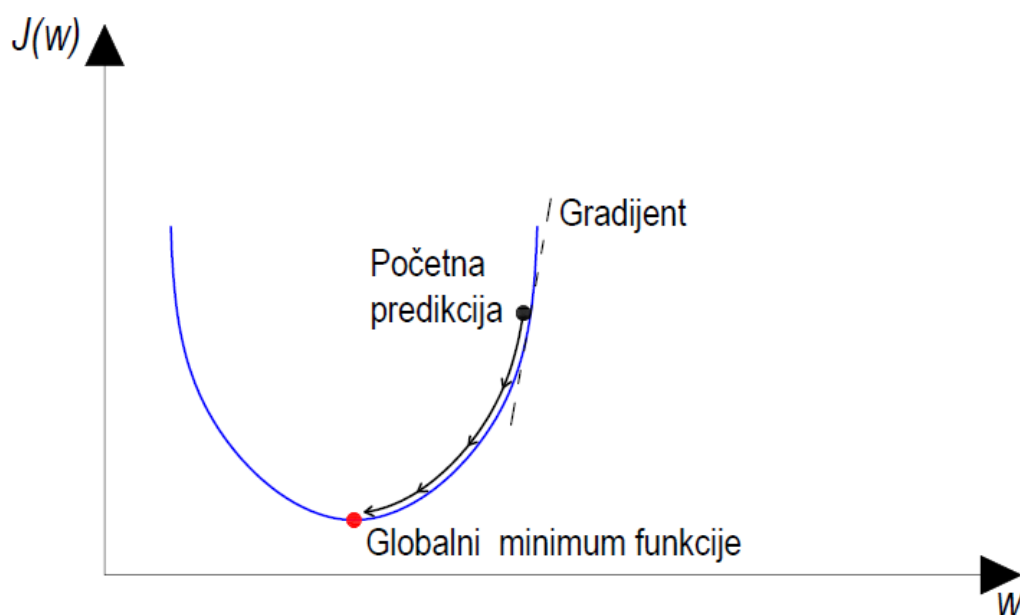
Slika 3. Generička „*feedforward*“ umjetna neuronska mreža s 4 ulazne jedinice, 2 izlazne jedinice i N skrivenih slojeva

1.2.1. Modeli strojnog učenja

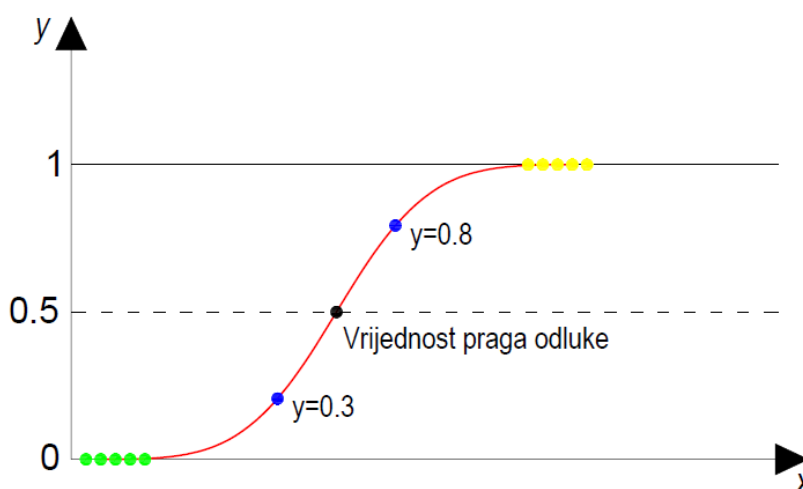
U nastavku su navedeni algoritmi koji su korišteni u razvijanju modela u ovoj disertaciji.

1.2.1.1. Logistička regresija

Logistička regresija (engl. *logistic regression*, LR) je diskriminativni klasificirajući model. U njoj se uzimaju prethodno označene vrijednosti kao ulazne varijable, kojima se potom u fazi treniranja dodijele određene težine i *bias term* (intercept), a na kraju se primijeni sigmoidna funkcija kako bi se dobila vjerojatnost nekog ishoda. Sigmoidna funkcija također se naziva logistička po čemu je ovaj model i dobio ime. Konačna izlazna varijabla dobiva se ovisno o postavljenom pragu odluke. LR se može koristiti za binarne ishode, ali i za više klasa što se naziva multinomijalnom logističkom regresijom. Ova podvrsta koristi *softmax* funkciju kako bi se izračunale vjerojatnosti. Težine su naučene iz skupa podataka za treniranje putem funkcije gubitka koja se mora svesti na minimum (**Slika 4.**). Sigmoidna funkcija ima više prednosti; ona uzima vrijednost i svrstava je u raspon od 0 do 1 što je upravo ono što je potrebno kod vjerojatnosti. Nadalje, gotovo je linearna oko središnje vrijednosti ali odstupajuće vrijednosti (engl. *outliers*) bivaju spljoštene na rubovima oko 0 i 1 (**Slika 5.**)(115).



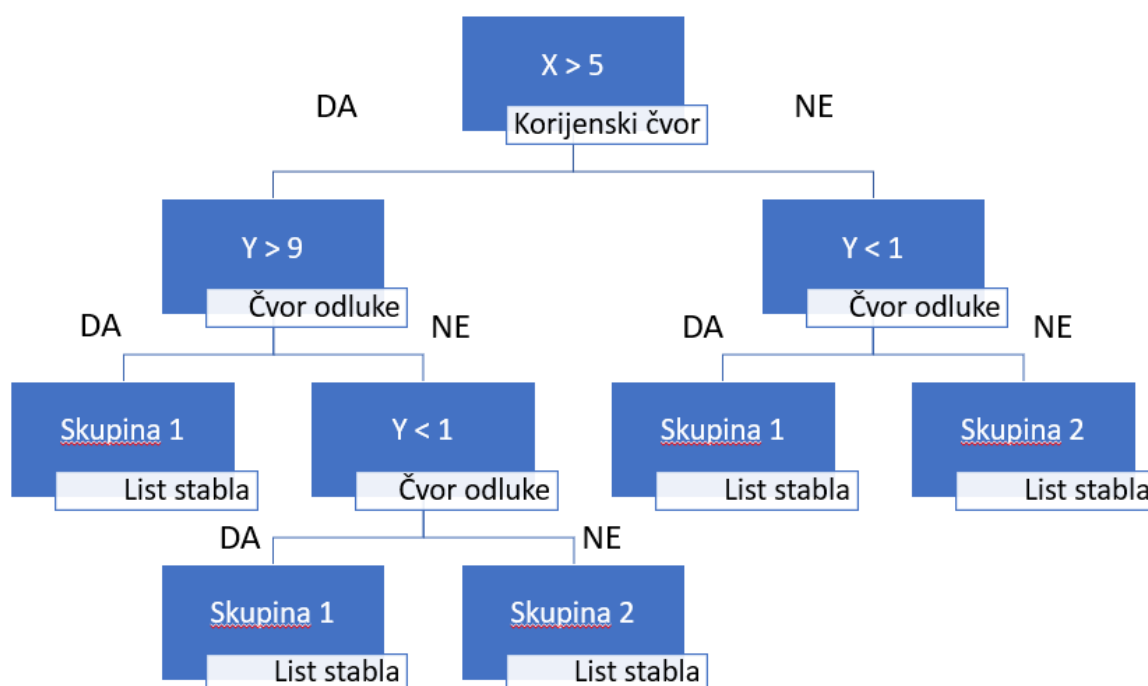
Slika 4. Slikovni prikaz kvadratne funkcije gubitka



Slika 5. Prikaz sigmoidne funkcije u logističkoj regresiji s postavljenim pragom odluke 0,5

1.2.1.2. Slučajne šume

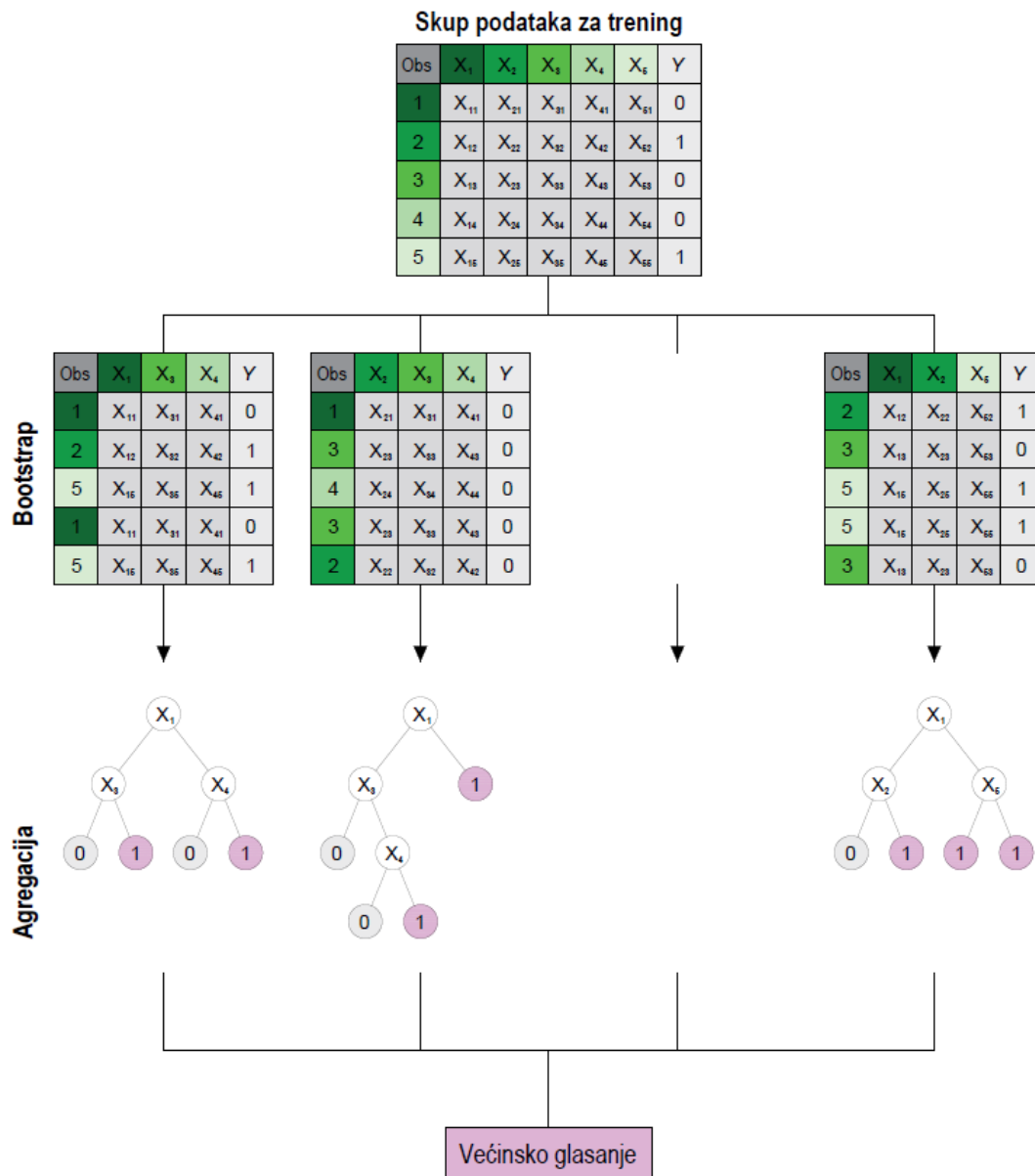
Kako bi se razumjele slučajne šume (engl. *random forest*, RF), potrebno je definirati stabla odluke jer čine osnovu RF-a. Ona predstavljaju algoritam za izračunavanje ishoda funkcije $f(x)$ putem niza uzastopnih *if/else* testova na ulaznoj varijabli X . Ishod svakog testa određuje sljedeći korak, sve dok se konačni rezultat ne utvrdi s potpunom sigurnošću. Svaki čvor (engl. *node*) predstavlja *if/else* funkciju, a ulazna varijabla prolazi kroz sve čvorove, dok posljednji, u kojem se odlučuje u koju će klasu varijabla biti razvrstana, naziva se list stabla (engl. *leaf node*). Učenje stabla odluke odnosi se na zadatak izrade stabla iz seta parova x i $f(x)$, koji odražava f ili približnu aproksimaciju. Kada je domena ulaznih podataka konačna, skup podataka može biti iscrpan, no često je riječ o uzorku iz veće, moguće beskonačne domene. U tom slučaju, cilj je pronaći stablo koje približno odgovara funkciji f na cijeloj domeni, a ne samo na uzorku podataka. Osim pronalaženja dobre aproksimacije, mogu biti postavljeni i dodatni kriteriji, poput pronalaska najjednostavnijeg stabla koje odražava funkciju. Izlazna varijabla za ulaznu varijablu x se naziva predikcija za x . Stabla odluke koja rade predikciju nominalnih varijabli nazivaju se klasifikacijska stabla, a ona koja rade predikciju numeričkih varijabli regresijska stabla. Njihova ključna prednost u kontekstu strojnog učenja jest interpretativnost – predikcije se temelje na jednostavnom nizu testova koji se lako tumače, čineći stabla odluka preglednim i razumljivim modelima (**Slika 6.**) (116, 117).



Slika 6. Primjer stabla odluke s 3 čvora odluke i 5 listova

Međutim, stabla odluke mogu pretjerano prilagoditi podatke ako se ne odrede odgovarajući kriteriji u procesu izrade stabla, što se naziva prenaučenos (engl. *overfitting*). To znači da je model previše dobro naučio predviđati vrijednosti na kojima je stvoren, pa ne generalizira dobro na novouvedenim vrijednostima (118). RF modeli su dizajnirani kako bi prevladali glavno ograničenje stabala odluke, a to je da je jedno pojedinačno stablo visoko ovisno o podacima i „nestabilno“. Taj problem je riješen tako što se stvori više stabala odluke uzimajući nasumično odabrane uzorke iz skupa podataka umjesto iz cijelog skupa. Predikcije iz stabala izrađenih iz uzoraka na koje je primijenjena *bootstrap* metoda potom se agregiraju kako bi se dobile preciznije procijenjene vrijednosti nego one iz jednog stabla: ovaj postupak poznat je kao „*bagging*“ (skraćeno od *bootstrap aggregating*). Osim randomizacije u uzorkovanju podataka, svako stablo odluke u slučajnim šumama izrađuje se nasumičnim uzorkovanjem podskupa prediktora na svakom grananju. Kao rezultat, stabla unutar šume postaju različitija jedno od drugog, čime se smanjuje pogreška u generalizaciji šume i poboljšava preciznost predikcije. Skup izlaznih varijabli svakog stabla na kraju se agregira za konačnu predikciju, čime se smanjuje varijabilnost, poboljšava ukupna točnost predikcije i smanjuje prenaučenos (118).

U slučajnim šumama važnost varijabli se procjenjuje kroz permutaciju varijabli i ponovno prilagođavanje modela. Svako stablo unutar šume ponovno se prilagodi tako što se nasumično permutiraju vrijednosti jedne od prediktorskih varijabli. Posljedično, ova nasumična permutacija narušava povezanost između prediktora i zavisne varijable (ciljna ili izlazna varijabla – ishod), tako da bi se točnost predikcije trebala smanjiti ako je prediktor važan u određenom stablu dok točnost ostaje ista ili čak raste ako prediktor nije važan u tom stablu. Uzimanje prosjeka vrijednosti u svim stablima za svaki prediktor daje procjenu relativne važnosti svakog prediktora (**Slika 7.**) (118).



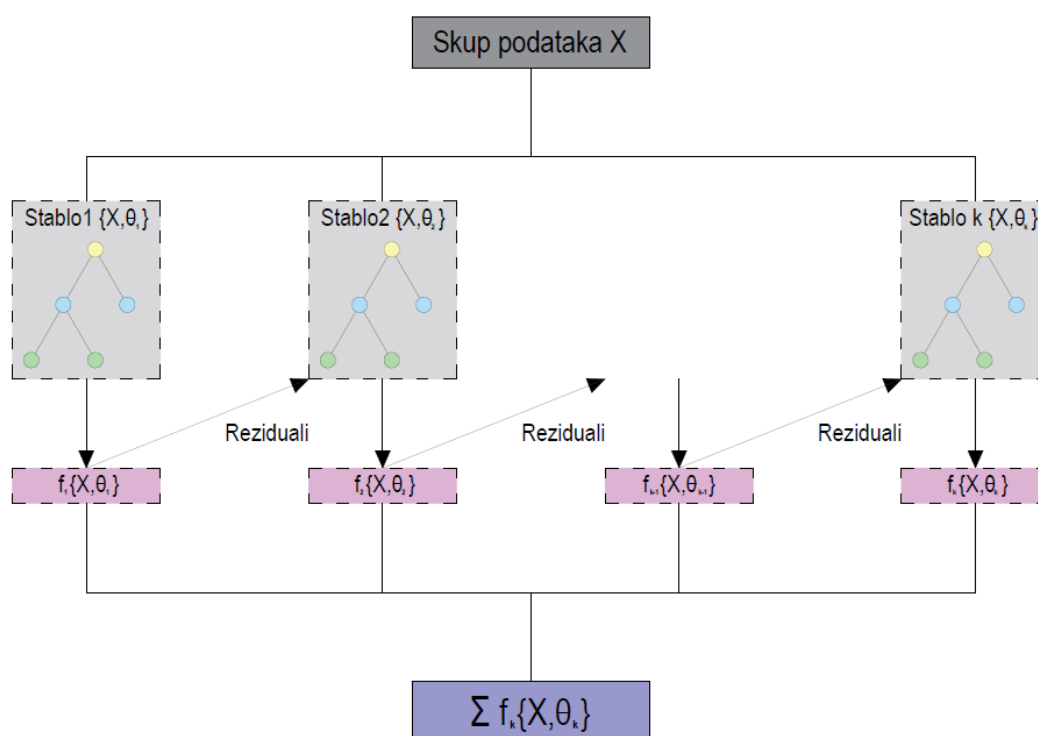
Slika 7. Shematski prikaz algoritma slučajnih šuma. (Preuzeto i prilagođeno prema: Montesinos López OA, Montesinos López A, Crossa J. Random forest for genomic prediction. U: Montesinos López OA, Montesinos López A, Crossa J, urednici. Multivariate statistical machine learning methods for genomic prediction. Cham: Springer; 2022. str. 633-81. Djelo otvorenog pristupa licencirano pod CC-BY 4.0 licencom)

1.2.1.3. eXtreme gradient boosting

Extreme gradient boosting (XGBoost) je metoda strojnog učenja temeljena na ansamblu, tj. agregaciji rezultata više pojedinačnih stabala odluke (117). Ono što ovu metodu razlikuje od slučajnih šuma jest to što se ovdje koristi metoda podizanja gradijenta (engl. *gradient boosting*). Ova metoda kombinira slabe klasifikatore (engl. *weak learners*), odnosno klasifikatore koji su nešto bolji od nasumičnosti, u jake klasifikatore (engl. *strong learners*) na

iterativan način. Cilj metode podizanja gradijenta je pronaći aproksimaciju funkcije $f(x)$ u kojoj su izlazne vrijednosti y , tako da se minimalizira funkcija gubitka $L(y, f(x))$. Ova metoda gradi aditivni model dodavanjem novih stabala koji ispravljaju greške prethodnih (**Slika 8.**)(119).

Nadalje, XGBoost je posebno dizajniran za brzinu i efikasnost. Ova se metoda fokusira isključivo na stabla odluke kao osnovne klasifikatore i uključuje različite tehnike regularizacije za kontrolu složenosti stabala. XGBoost također integrira i metode randomizacije kako bi se smanjila prenaučenosť (119). Kako bi se maksimizirala preciznost XGBoost-a, radi se na optimizaciji hiperparametara kao što su broj stabala odluke, najveća „dubina“ stabla odluke i drugi. U samom procesu optimizacije hiperparametara tipično se koristi proces unakrsne validacije. Tu se radi o nasumičnoj podjeli skupa podataka za treniranje kako bi se stvorio mali validacijski set koji se koristi za mjerenje performansi. Na ovaj način se može kvantificirati učinak promjene hiperparametara i odabrati optimalna vrijednost (117).



Slika 8. Shematski prikaz *eXtreme Gradient Boosting* algoritma. *reziduali* – razlike između promatranih i predviđenih vrijednosti. (Preuzeto i prilagođeno prema: Guo R, Zhao Z, Wang T, Liu G, Zhao J, Gao D. Degradation state recognition of piston pump based on iceemdan and xgboost. Appl Sci. 2020;10:6593. Djelo otvorenog pristupa licencirano pod CC-BY 4.0 licencom)

1.2.2. Validacija modela strojnog učenja

Validacija je značajan korak u razvoju modela strojnog učenja pomoću kojeg procjenjujemo performanse i generaliziranost modela. Naučenost modela (engl. *model fit*) je pod utjecajem ravnoteže između pristranosti (bias) i varijance, a oboje su povezani s kompleksnošću modela. Manje kompleksni modeli imaju tendenciju slabije naučenosti (engl. *underfit*) što znači da imaju visoku pristranost i nisku varijancu te ne mogu precizno uhvatiti složene odnose između ulaznih i izlaznih varijabli, što u konačnici rezultira slabim performansama kako na skupu podataka za trening tako i na skupu podataka za testiranje. Suprotno tome, kompleksniji modeli skloni su prekomjernoj naučenosti što rezultira odličnim performansama na skupu podataka za trening, ali visokim stupnjem pogreške na skupu podataka za test zbog niske pristranosti i visoke varijance. Optimalan razvoj modela zahtijeva kompromis između pristranosti i varijance kako bi se spriječila i slaba i prekomjerna naučenost. Tehnike poput *bagginga* i *boostinga* često se koriste u vaganju između pristranosti i varijance (120).

Postoji više pristupa za validaciju i evaluaciju modela kao što su *holdout* validacija, unakrsna validacija i njezine varijante uključujući ugniježdenu unakrsnu validaciju (121). U vanjskoj petlji ugniježdene unakrsne validacije greška generalizacije procjenjuje se korištenjem prosjeka grešaka dobivenih na skupovima podataka za test iz vanjske petlje, dok unutarnja petlja uzima skup podataka za trening i test odabran od vanjske petlje te se na njima odabiru najbolji hiperparametri ovisno o performansama treniranog modela na validacijskom skupu.

Validiranje podataka koristeći podatke drugog centra koji nije bio uključen u inicijalnu izradu modela naziva se vanjskom (eksternom) validacijom, dok se korištenje podataka centra uključenog u razvoj modela naziva unutarnjom (internom) validacijom (120).

Za brojčano izražavanje performansi predikcije koristi se metoda ovisno o izlaznoj varijabli, pa su ispod navedene neke metode za binarne i ordinalne varijable.

Matrica zabune (engl. *confusion matrix*) je alat kojim se može vizualizirati performanse modela uspoređujući prave oznake klasa s onima koje je model predvidio. Iz ove tablice se mogu izračunati vrijednosti kao što su stvarno pozitivni (engl. *true positives*, TP), stvarno negativni (engl. *true negatives*, TN), lažno pozitivni (engl. *false positives*, FP) i lažno negativni (engl. *false negatives*, FN) slučajevi (**Slika 9.**) (121).

		Predviđena kategorija		
		1	2	3
Stvarna kategorija	1	684	24	12
	2	20	593	4
	3	16	12	2024

Slika 9. matrica zabune na primjeru 3 kategorije

Nadalje, iz ovih vrijednosti može se izračunati osjetljivost, tj. odziv (engl. *recall*) modela kao $TP/(TP + FN)$ što označava sposobnost modela da prepozna stvarno pozitivne slučajeve. Nadalje, može se izračunati preciznost koja je definirana kao omjer stvarno pozitivnih slučajeva među svim pozitivnim predikcijama $TP/(TP + FP)$, te specifičnost kojom se procjenjuje sposobnost modela da točno klasificira negativne slučajeve $TN/(TN + FP)$. Proporcija točno klasificiranih slučajeva (engl. *proportion of cases correctly classified*) mjera je sveukupne preciznosti modela koja se računa iz omjera $(TP + TN)/ukupan broj slučajeva$. Međutim, potrebno je odabrati prikladan način za prikaz performansi modela, što na primjer može ovisiti i o balansiranoj setu podataka ili broju klasa. Jedna od popularnijih metrika za binarne podatke je površina ispod ROC krivulje (engl. *area under the curve*, AUC) koja daje veći prioritet pozitivnim predikcijama u odnosu na negativne. Ova krivulja je definirana kao graf na kojem je na x-osi vrijednost 1-specifičnost, tj. stopa lažno pozitivnih slučajeva, dok je na y-osi osjetljivost modela. Ovaj pristup je učinkovit za procjenu kvalitete ili performansi dijagnostičkih testova i koristi se u strojnom učenju za procjenu točnosti predikcija algoritma (121).

Matthewsov koeficijent korelacije jedan je od pokazatelja koji povezuje sve četiri mjere u matrici zabune. U strojnom učenju koristi se za procjenu kvalitete binarnih klasifikacija i može biti koristan kada postoji značajna neuravnoteženost između klasa. Ovaj pokazatelj zapravo je specifičan oblik Pearsonovog koeficijenta korelacije u binarnom slučaju i postiže

visoke vrijednosti samo ako se postižu dobri rezultati u svim unosima matrice zabune (121, 122).

F1 ocjena je metoda prikaza performansi u klasifikacijama koja kombinira preciznost i odziv koristeći njihovu harmonijsku sredinu. Definira se kao: $F1 = 2x (\text{preciznost} * \text{odziv}) / (\text{preciznost} + \text{odziv})$. Stoga je F1 score koristan kao mjera kojom se mogu uspoređivati klasifikacijske metode. Važno je dobro odabrati situaciju u kojoj se može koristiti ova metoda. Međutim, postoje i manjkavosti ove metode. Gledajući formulu, F1 ocjena ne uzima u obzir stvarno negativne rezultate; različiti parovi preciznosti i odziva mogu dati istu F1 ocjenu. Također, ova metrika nije simetrična između klasa, odnosno ovisi o tome koja je klasa definirana kao pozitivna, a koja kao negativna. Na primjeru gdje imamo vrlo neuravnotežene klase, stopa greške može biti pod utjecajem velikog broja nebitnih ispravnih klasifikacija unutar klase koja nam nije od važnosti i imali bi niske F1 vrijednosti. Nasuprot tome, u slučaju velike pozitivne klase i klasifikatora pristranog prema toj većinskoj klasi, vrijednosti F1 ocjene bile bi povišene (122, 123).

1.2.3. Objašnjivost modela strojnog učenja

Objašnjivost u strojnom učenju ima za cilj učiniti složene modele razumljivijima, pri čemu su Shapleyeve vrijednosti ključna tehnika. Podrijetlom iz teorije igara, Shapleyeve vrijednosti pravedno raspodjeljuju doprinos svake značajke u izlaznoj varijabli stvorenog modela. Ova metoda razmatra sve moguće kombinacije značajki, računajući prosječne doprinose kako bi pružila sveobuhvatno razumijevanje važnosti značajki. Shapleyeve vrijednosti široko se primjenjuju u različitim zadacima strojnog učenja, uključujući objašnjavanje predikcija modela dubokog učenja. Međutim, njihov izračun može biti računalno zahtjevan za modele s mnogo značajki, što je dovelo do razvoja aproksimacijskih metoda. Iako vrijedne, Shapleyeve vrijednosti imaju ograničenja, posebice u kontekstu visokodimenzionalnih prostora ili složenih međuodnosa značajki (124, 125).

1.2.4. Strojno učenje u medicini i abdominalnoj kirurgiji

Strojno učenje se sve više primjenjuje u zdravstvu kako bi se unaprijedio proces dijagnostike, liječenja i poboljšali ishodi pacijenata. Analizirajući zdravstvene kartone, ovi modeli bi mogli predvidjeti dijagnoze, identificirati pacijente s visokim rizikom od razvoja određene bolesti i optimizirati načine liječenja. Osim toga, ML bi se mogao koristiti i u radiološkoj dijagnostici. Razvijeni su modeli za prepoznavanje tumora, prijeloma i drugih

lezija. Nadalje, opisane su i primjene ML-a u razvijanju individualnog plana liječenja, kao i primjena u genetičkom inženjeringu i genomici (126).

Postoji širok raspon potencijalne koristi AI-a u abdominalnoj kirurgiji. Automatiziranom analizom svih predoperativnih podataka mogli bi se prepoznati individualni rizici pacijenta u operacijskom planiranju te pružiti vrijedni prediktori za postoperativnu skrb. Kirurzi bi mogli poboljšati svoj proces donošenja odluka tijekom operacije, temeljen na analizi napretka operacije u stvarnom vremenu, koja bi koristila elektroničke medicinske zapise kombinirane s video zapisom operacije, vitalnim znakovima, praćenjem pokreta instrumenata i manualnih pokreta te korištenjem podataka o potrošnji elektrokirurške energije. Sve zajedno bi moglo dovesti do predikcije i posljedičnog izbjegavanja štetnih događaja. Integracija pre-, intra- i postoperativnih podataka bi mogla pomoći u praćenju oporavka i predviđanju komplikacija. Osim toga, potencijalna upotreba AI-a u kirurgiji uključuje automatiziranu robotsku kirurgiju, pomoć u procesu odlučivanja u stvarnom vremenu, implementaciju u edukaciji kirurga te procjenu izvedbe zahvata (127, 128).

Nadalje, kreirani AI modeli koji pružaju automatizirano strukturiranje i analizu snimljenih video zapisa prema ključnim intraoperativnim događanjima mogu pružiti pregled proceduralnih koraka, sigurnosnih prekretnica (engl. *safety milestones*), stupnjevanje težine bolesti te pregled intraoperativnih događaja (128). Opisano je korištenje AI modela u laparoskopskoj kolecistektomiji, laparoskopskoj i robotskoj sleeve gastrektomiji, laparoskopskoj rektomektomiji, laparoskopskoj kirurgiji kolona, laparoskopskom totalnom ekstrapéritonealnom popravku preponske kile i laparoskopskoj apendektomiji (128-130).

1.2.5. Modeli strojnog učenja u akutnom apendicitisu

Korištenje AI-a u dijagnostici i liječenju akutnog apendicitisa još je uvijek u povojima. Provedene su studije koje su se bavile dijagnozom, diferencijacijom težine bolesti, načinom liječenja i predikcijom postoperativnih komplikacija. Modeli su kreirani korištenjem demografskih, anamnestičkih, kliničkih, laboratorijskih, radioloških i patohistoloških podataka (131). Ovisno o svrsi razvoja modela možemo ih podijeliti u četiri skupine: 1) modeli za korištenje u dijagnozi apendicitisa, 2) modeli za korištenje u predviđanju liječenja apendicitisa, 3) modeli za korištenje u predviđanju postoperativnih komplikacija apendektomije i 4) modeli za korištenje u operacijskom liječenju apendicitisa.

1.2.5.1. Modeli za korištenje u dijagnozi apendicitisa

Do sada provedena istraživanja u dijagnozi apendicitisa korištena su u samoj dijagnozi u užem smislu kao i u diferencijaciji težine bolesti u vidu diferencijacije nekomplikiranih i komplikiranih oblika apendicitisa. U tome su se koristili raznim ML i DL modelima. Pri tome su uglavnom korišteni demografski i laboratorijski podaci. Tako su Aydin i sur., Mijwil i sur., Akbulut i sur. u svojim istraživanjima razvijali ML modele bazirane na demografskim, kliničkim i laboratorijskim nalazima u svrhu dijagnoze i diferencijacije težine bolesti (132-134). Ghareeb i sur., Reismann i sur. koristili su daljnje dijagnostičke modalitete kao što su ultrazvuk i klinički sustavi bodovanja (135, 136).

Zatim, provedena su istraživanja koja su koristila DL modele. Park i sur., Prabhudesai i sur., te Yoldas i sur. razvili su DL modele bazirane na kliničkim, laboratorijskim nalazima kao i kliničkim sustavima bodovanja (137-139). Također, provedena su istraživanja koja su koristila i ML i DL za dijagnozu apendicitisa. Sakai i sur., Akmese i sur. razvili su modele bazirane na kliničkim i laboratorijskim nalazima, dok su Phan-Mai i sur. dodatno koristili UZV, a Hsieh i sur. koristili su klinički sustav bodovanja (140-143).

Osim navedenih, Stiel i sur. su koristili ML modele u razvijanju sustava bodovanja, a Rajpurkar i sur. su razvili 3D konvolucijsku neuronsku mrežu (CNN) za automatiziranu detekciju apendicitisa na CT-u abdomena (144, 145). U drugom istraživanju baziranom na ekspresiji gena Reismann i sur. uspjeli su povezati set od četiri gena kao diskriminatorni u predviđanju flegmonoznog ili gangrenoznog apendicitisa (146).

1.2.5.2. Modeli za korištenje u predviđanju liječenja apendicitisa

Osim korištenja u dijagnozi, razvijeni su modeli i za predviđanje PHD nalaza kako bi se usmjerio tijek liječenja na konzervativno ili kirurško. Kang i sur. razvili su model za predviđanje patološkog tipa akutnog apendicitisa koji je bio baziran na kliničkim pokazateljima i biomarkerima, od kojih neki nisu u standardnoj uporabi u dijagnostici akutnog apendicitisa, a smatraju da pomaže u planiranju načina liječenja i uporabi antibiotika (147). Marcinkevičs je sa suradnicima u dva odvojena istraživanja razvio modele bazirane na predikciji težine PHD nalaza, koji bi se mogli koristiti kao potpora u odlučivanju načina liječenja (148, 149). U prvom istraživanju razvijen je ML model za predviđanje dijagnoze, načina liječenja i težine bolesti koji je temeljen na anamnestičkim, kliničkim i laboratorijskim pokazateljima te nalazom UZV-a (148). U drugom istraživanju stvoreni su modeli bazirani na svemu prethodno spomenutom, uz razliku što su modeli trenirani na samim slikama UZV-a (149). Liang i sur. razvili su modele

temeljene na dubokom učenju i radiomici koristeći kliničke i CT značajke u diferencijaciji nekomplikiranog od kompliciranog apendicitisa. Dobiveni model pokazao je dobre karakteristike koje bi mogle biti koristan dodatak trenutnom dijagnostičkom procesu te pomoći prilagođavanju konzervativnog liječenja apendicitisa (150).

1.2.5.3. Modeli za korištenje u predviđanju postoperativnih komplikacija

Za razliku od korištenja u dijagnostici, manje je provedenih istraživanja o ulozi ML-a u predviđanju postoperativnih komplikacija, među kojima prednjači pojava intraabdominalnih apscesnih kolekcija. Osim toga, istraživana je i pojava drugih postoperativnih komplikacija kao što je sepsa, potreba za intenzivnim liječenjem, reoperacija, infekcija kirurške rane, duljina postoperativnog boravka u bolnici (131).

1.2.5.4. AI modeli za korištenje u operacijskom liječenju apendicitisa

Dayan i sur. proveli su validacijsku studiju i studiju o pouzdanosti o implementaciji AI-baziranih modela temeljenih na računalnom vidu (engl. *computer vision*) u laparoskopskoj apendektomiji. Stariji kirurzi složili su se da težina slučaja, procijenjena putem AI-a, bila točna u 76,9% do 94,4% slučajeva. Nadalje, stariji kirurzi procijenili su pridržavanje specijalizanata sigurnosnim protokolima kao visoko, tako što je procijenjeno da je napravljen kompletni kritični pogled sigurnosti (engl. *critical view of safety*, CVS) u 92% i 91,4% slučajeva, dok je AI model točno procijenio dovršenost kompletnog CVS-a u 84,4% slučajeva. Međutim, kada je model procijenio da je napravljen kompletni CVS, kirurzi su se složili u 99,8% i 96% slučajeva. Autori navode da se model doima manje točan u procjeni djelomično dovršenog CVS-a, što pripisuju dizajnu modela koji procjenjuje CVS u trenutku prije ekscizije crvuljka i ne uzima u obzir ranija razdoblja dovršenosti CVS-a (128).

2. CILJEVI I HIPOTEZE ISTRAŽIVANJA

Glavni cilj istraživanja:

Razviti i validirati modele strojnog učenja za predikciju akutnog apendicitisa u pedijatrijskoj i odrasloj populaciji koristeći kliničke i laboratorijske parametre dostupne pri prijemu u bolnicu.

Sporedni ciljevi istraživanja su:

1. Razviti i optimizirati modele strojnog učenja za detekciju apendicitisa u pedijatrijskoj i odrasloj populaciji s prioritetom postizanja visoke osjetljivosti
2. Procijeniti performanse razvijenih modela za razlikovanje pozitivnih od negativnih patohistoloških nalaza apendektomije u pedijatrijskoj i odrasloj populaciji
3. Procijeniti performanse razvijenih modela za predikciju kompliciranih oblika apendicitisa u odnosu na nekomplicirane oblike u pedijatrijskoj i odrasloj populaciji
4. Koristeći Shapley vrijednosti, identificirati najvažnije kliničke i laboratorijske prediktore prema razvijenim modelima
5. Usporediti performanse najboljeg modela u pedijatrijskoj populaciji s AIR sustavom bodovanja
6. Provesti *bootstrap* validaciju za procjenu robusnosti i stabilnosti razvijenih modela u odrasloj populaciji pri različitim uvjetima uzorkovanja

Hipoteze:

1. Modeli strojnog učenja temeljeni na kliničkim i laboratorijskim parametrima će postići visoku osjetljivost u predikciji akutnog apendicitisa uz zadovoljavajuću specifičnost u pedijatrijskoj i odrasloj populaciji.
2. Modeli strojnog učenja temeljeni na kliničkim i laboratorijskim parametrima će postići visoku osjetljivost u predikciji kompliciranog oblika akutnog apendicitisa uz zadovoljavajuću specifičnost u pedijatrijskoj i odrasloj populaciji.
3. Modeli strojnog učenja će pokazati značajno bolje performanse u odnosu na AIR sustav bodovanja
4. *Bootstrap* validacija će potvrditi robusnost modela s uskim intervalima pouzdanosti za ključne metrike performansi

3. ISPITANICI I POSTUPCI

3.1. Ispitanici

Uzorak se sastoji od bolesnika koji su operirani pod kliničkom slikom akutnog apendicitisa u KBC Split. Za bolesnike hospitalizirane u Klinici za dječju kirurgiju prikupljeni su podaci u razdoblju od 01/2019 do 06/2023, a za bolesnike koji su bili hospitalizirani u Klinici za kirurgiju prikupljeni su podaci od 01/2020 do 06/2024.

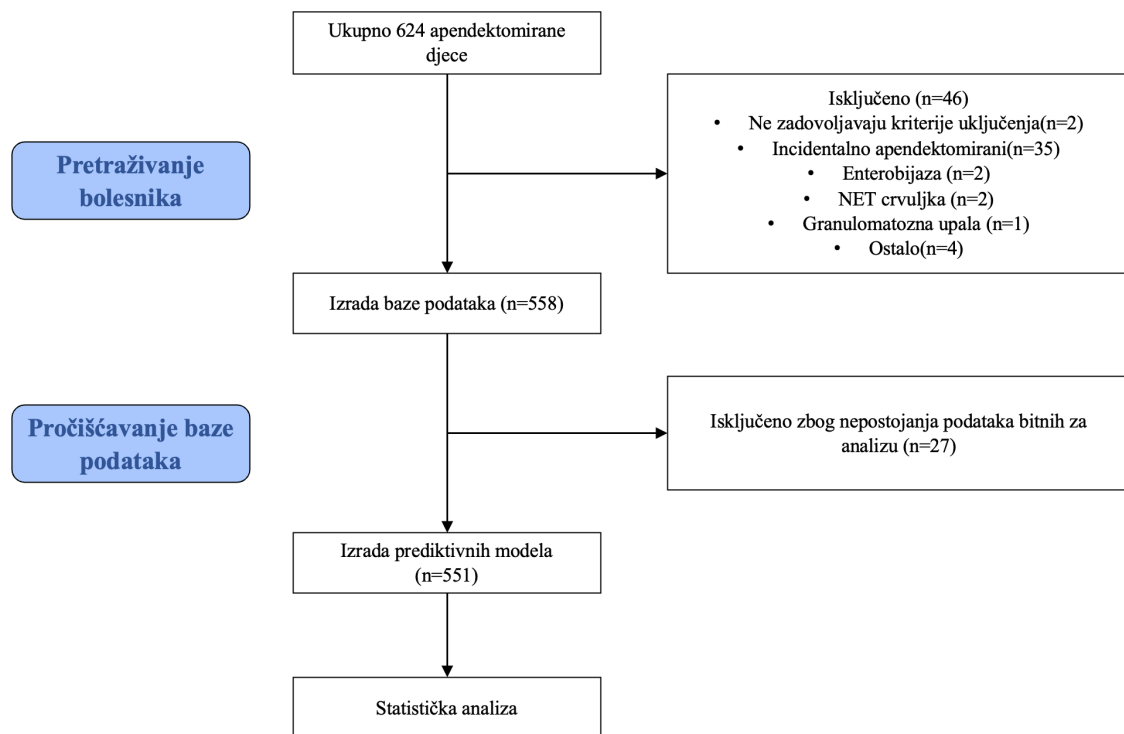
Podaci su retrospektivno prikupljeni iz bolničkog informacijskog sustava (BIS) tako što su pretraživani svi bolesnici kojima je učinjena klasična apendektomija ili laparoskopska apendektomija, uz napomenu da nisu prikupljeni podaci o bolesnicima kojima je isti postupak učinjen incidentalno. Također nisu prikupljeni podaci o bolesnicima kojima nalaz PHD-a nije bio dostupan ili je dijagnoza bila nešto drugo osim akutnog apendicitisa ili inoentnog crvuljka (npr. enterobijaza, NET crvuljka). Primijenjeni su sljedeći kriteriji isključenja u pedijatrijskoj populaciji: dob <17 godina za analizu u pedijatrijskoj populaciji, prisutnost bilo kojeg značajnog komorbiditeta koji bi mogao utjecati na preoperativne laboratorijske vrijednosti te BMI ≥ 35 kod djece. Nakon primjene navedenih kriterija, baza podataka se sačinjavala od 615 ispitanika, dok se baza podataka sačinjavala od 1167 ispitanika u odrasloj populaciji.

Istraživanje je odobreno od strane Etičkog povjerenstva KBC-a Split (Klasa: 500-03/23-01/169; Ur. br. 2181-147/01/06/LJ.Z.-23-02; 21.7.2023. zatim Klasa: 500-03/23-01/169; Ur. br. 2181-147/01/06/LJ.Z.-23-04; 21.7.2023. te Klasa: 500-03/23-01/169, Ur.broj: 2181-147/01-06/LJ.Z.-24-06; 24.10.2024.) te je provedeno u skladu s etičkim principima Helsinške deklaracije iz 2013. godine. Svi osobni podaci su bili anonimizirani.

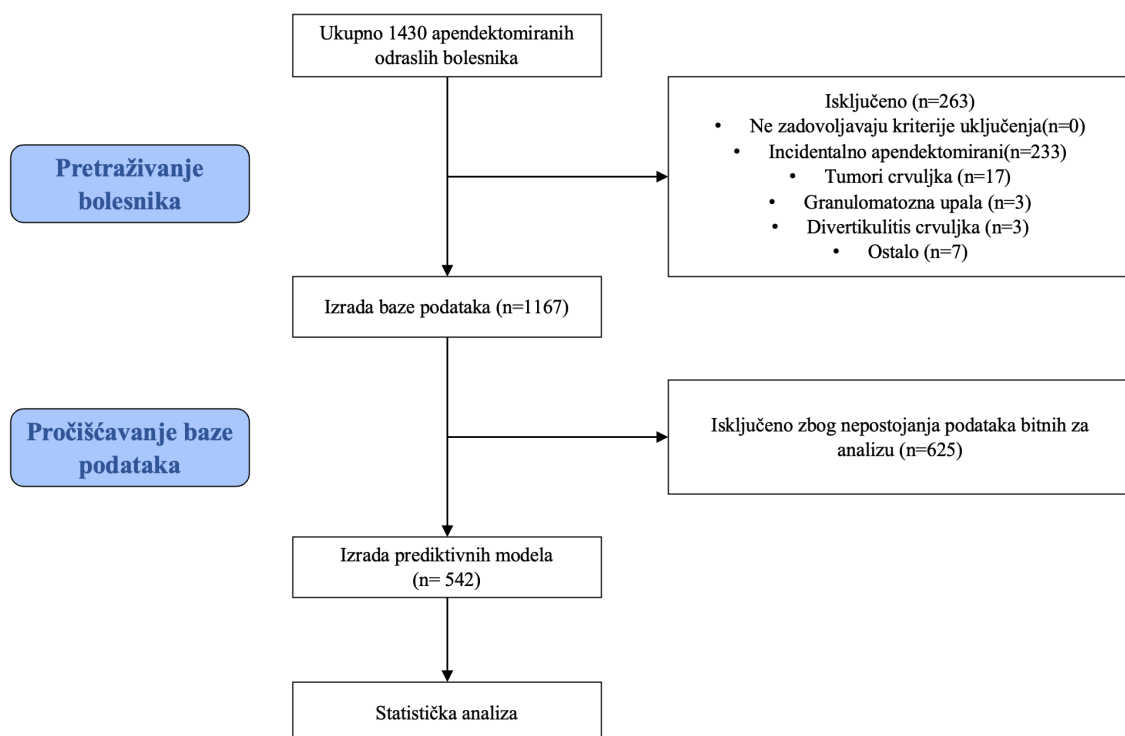
3.2. Postupci

Izrađena baza dodatno je provjerena za pogrešno upisane podatke, pri čemu su korigirane ili izbačene stavke s ekstremno visokim vrijednostima koje se ne mogu pronaći u stvarnim nalazima. Također su standardizirani formati različitih varijabli. Svi patohistološki nalazi su pojedinačno pregledani i kodirani u tri skupine ovisno o nalazu: inoentni, nekomplikirani (kataralni ili flegmonozni) i komplikirani (gangrenozni ili gangrenozni perforirani). Određene inicijalno prikupljene vrijednosti su isključene iz daljnje obrade ako su nedostajale u više od 30% bolesnika. Također, vrijednosti su isključene iz daljnje obrade ako koreliraju s nekim drugim izabranim vrijednostima (npr. hemoglobin, hematokrit, MCH i slično). Potom se iz daljnje analize isključuje dio bolesnika kod kojih nedostaje više od dvije vrijednosti za koje je ustanovljeno da su od visokog značaja za analizu bazirano na prethodnim

istraživanjima (leukociti, postotak neutrofila, trombociti, C-reaktivni protein i natrij) (Slika 10. i 11.).



Slika 10. Hodogram prikupljanja podataka u pedijatrijskoj populaciji

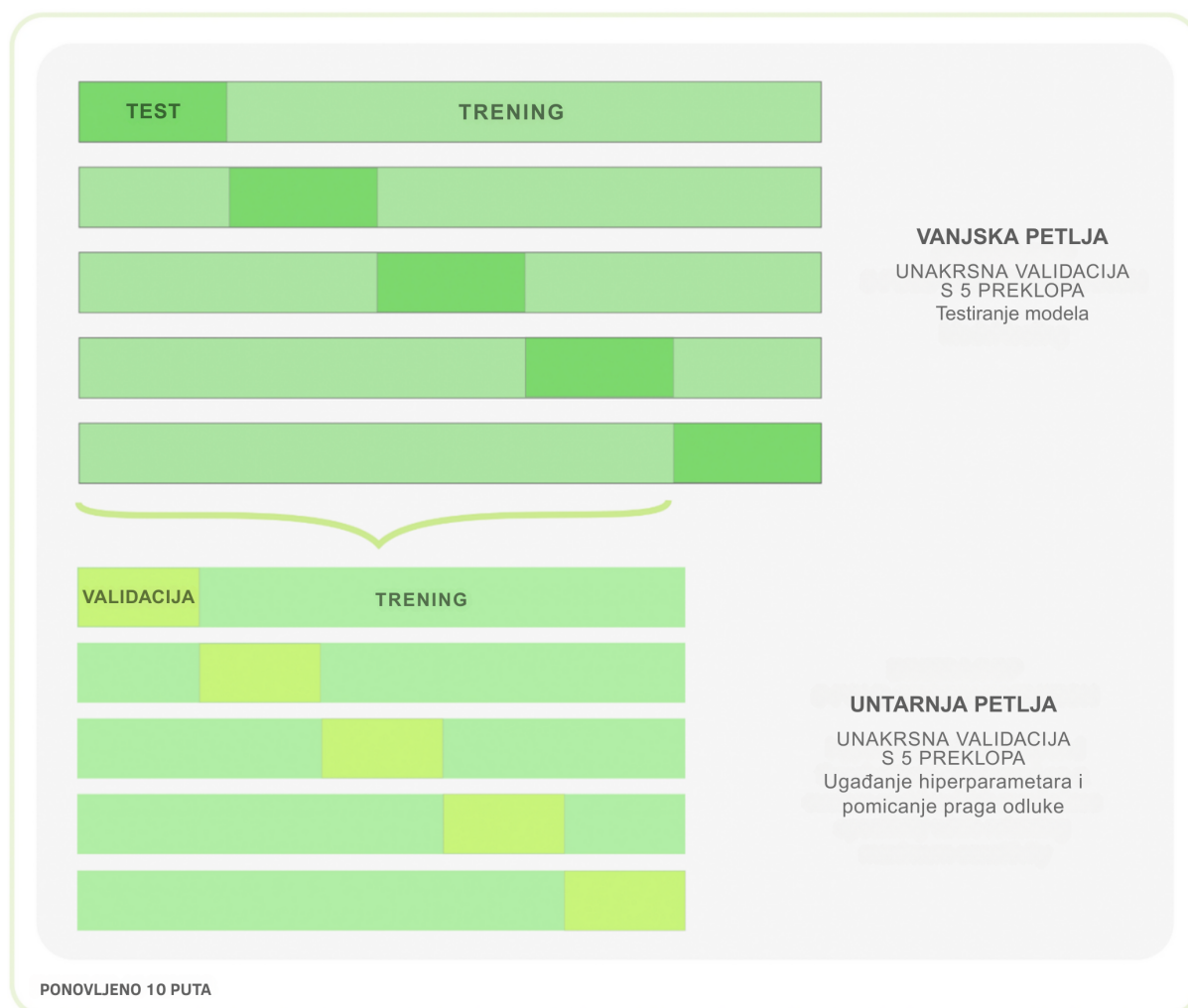


Slika 11. Hodogram prikupljanja podataka u odrasloj populaciji

3.2.1. Izrada modela i optimizacija

Izrađena su tri različita modela („Slučajne šume“, XGBoost i logistička regresija). U pedijatrijskoj populaciji potonja je izabrana kao osnovni model za usporedbu, dok su druga dva izabrana zbog svoje sposobnosti obrade tabličnih podataka i neuravnoteženih skupova podataka. A potom su isti modeli izrađeni i evaluirani u odrasloj populaciji.

Za treniranje i validaciju modela korišten je pristup unakrsne validacije s petostrukom unutarnjom i vanjskom ugniježđenom validacijom, ponovljenom 10 puta. Također je napravljena stratifikacija po ciljnoj varijabli zbog nebalansiranog skupa podataka, tako da su klase ciljne varijable prisutne u svakoj podjeli. Cijeli skup podataka dijeli se na 5 vanjskih preklopa, u svakoj vanjskoj iteraciji 4 preklopa tvore trening skup, a preostali preklop služi kao neovisni testni skup. Kao i u vanjskoj petlji, u svakoj unutarnjoj petlji 4 preklopa tvore skup za trening, a preostali preklop tvori unutarnji test, odnosno validacijski set. U unutarnjoj unakrsnoj validaciji se vrši optimizacija hiperparametara i pomicanja praga odluke, potom se te vrijednosti kao fiksne primjenjuju na neovisni vanjski testni set te iteracije. Na ovaj način se optimizacija i kalibracija odvijaju isključivo na podacima koji ne uključuju vanjski set, dok izvještavanje performansi proizlazi iz predikcija na pripadajućim vanjskim testnim setovima određene iteracije čime se dobiva nepristrana procjena modela (**Slika 12.**).



Slika 12. Prikaz postupka izrade modela

Određene vrijednosti koje su nedostajale imputirane su koristeći *Bagged Trees* algoritam uz pomoć „step_impute_bag“ funkcije u „recipes“ paketu unutar R programskog jezika. Kako bi se spriječilo curenje podataka, imputacija u validacijskom i test skupu je izvršena isključivo bazirano prema procjenama izračunatih iz njihovih pripadajućih trening skupova.

Izvršeno je optimiziranje hiperparametara modela. U pedijatrijskoj populaciji za slučajne šume optimizirani su *mtry* koji definira broj prediktora uzorkovanih na svakom grananju stabla i *min_n* koji postavlja minimalnu veličinu čvora potrebnu za daljnje grananje, čime se kontrolira dubina i složenost individualnih stabala i tako ograničava prekomjernu dubinu i prenaučenos modela. Za XGBoost optimizacija je obuhvatila: *learn_rate* (brzina učenja), koji modulira doprinos svakog uzastopnog stabla i stabilizira konvergenciju; *loss_reduction* (*gamma*), koji uvjetuje minimalnu redukciju gubitka za uspostavu novog razdvajanja i time penalizira nepotrebnu složenost; *trees* za ukupan broj stabala, kojim se

kontrolira kapacitet ansambla; *sample_size (subsample)*, koji određuje udio podataka korišten pri izgradnji pojedinog stabla i uvodi stohastičnost radi bolje generalizacije; *tree_depth* za maksimalnu dubinu stabala, kojim se izravno ograničava složenost pojedinih stabala. Ovakva konfiguracija hiperparametara u pedijatrijskom skupu uravnotežuje potrebu za visokom osjetljivošću s kontrolom složenosti modela, osiguravajući robusnost na ograničenim veličinama uzorka i potencijalnoj heterogenosti signala.

U odrasloj populaciji, za logističku regresiju optimizirani su sljedeći hiperparametri: 1) *penalty* za kontrolu snage regularizacije koja sprječava preučavanje modela, i 2) *mixture* za optimizaciju *elastic-net* miješanja između *ridge* regularizacije (L2) koja zadržava sve značajke uz smanjene koeficijente i *lasso* regularizacije (L1) koja omogućava selekciju značajki postavljanjem koeficijenata na nulu. Za slučajne šume, kao i u pedijatrijskoj populaciji optimizirani su *mtry* i *min_n* kao i *trees* koji određuje ukupan broj stabala u ansamblu. Za XGBoost algoritam, kao i u pedijatrijskoj populaciji optimizirani su *learn_rate*, *loss_reduction (gamma)*, *trees*, *sample_size (subsample)*, *tree_depth* te su dodatno optimizirani: 1) *min_n* za minimalnu veličinu čvora, 2) *mtry* za broj prediktora uzorkovanih u svakom čvoru, i 3) *stop_iter* za broj iteracija ranog zaustavljanja kada se performanse ne poboljšavaju, što sprječava preučavanje i poboljšava generalizaciju modela.

Pomak praga na krivulji odnosa specifičnosti i osjetljivosti klasifikatora (*engl. receiver operating characteristics, ROC*) korišten je za rješavanje problema neuravnoteženosti podataka. Budući da je u pedijatrijskoj populaciji naš cilj bio izgraditi model koji poboljšava detekciju negativnih slučajeva bez žrtvovanja sposobnosti detekcije pozitivnih slučajeva, odlučili smo prilagoditi hiperparametre modela i izvršiti pomak praga na temelju prilagođene metrike koja locira točku najviše specifičnosti na ROC krivulji, za koju je osjetljivost još uvijek 1 (*maksimum (specifičnost/osjetljivost) = 1*). Praga odluke primijenjen na testni skup određen je kao drugi najniži prag dobiven unutarnjom unakrsnom validacijom. U smislu zadržavanja visoke osjetljivosti na testnom skupu, ovo je konzervativniji pristup u usporedbi s uzimanjem srednje vrijednosti praga. Drugim riječima, postavili smo ograničenje da model mora biti 100% točan u dijagnosticiranju stvarno pozitivnih apendicitisa kod pacijenata u skupu trening podataka. Na taj način možemo biti što sigurniji da, kada model daje negativnu dijagnozu, ona nije lažno negativna, već stvarno negativna. Srednja vrijednost i standardna pogreška ciljanih metrika dobivene su prosjekom rezultata iz 50 vanjskih skupova testnih podataka (vanjska unakrsna validacija s pet odjeljaka ponovljena 10 puta). Najbolji model odabran je na temelju srednje vrijednosti specifičnosti uz zadržavanje maksimalne osjetljivosti.

Nadalje, za odraslu populaciju također je napravljena prilagođena metrika koja određuje točku na ROC krivulji koju smo postavili kao najveću specifičnost za određene pragove osjetljivosti kako bismo bolje procijenili ponašanje modela uzimajući u obzir reducirani broj varijabli za kreiranje modela i kako bi napravili usporediv model s onim pedijatrijskoj. Prag osjetljivosti u odrasloj populaciji primijenjen na testni skup određen je kombiniranjem svih predviđanja iz unutarnjih validacijskih skupova unutar svakog vanjskog preklopa te direktnom selekcijom pragova koji maksimiziraju specifičnost za tri različite razine osjetljivosti (0,98, 0,99 i <0,995). Ovaj pristup osigurava da se pragovi biraju na temelju sveukupnih unutarnjih validacijskih rezultata umjesto agregacije preko preklopa, što predstavlja robusniji i statistički stabilniji pristup u usporedbi s uzimanjem medijana ili drugih mjera centralne tendencije. Drugim riječima, postavili smo ograničenja koja omogućavaju modelu održavanje visoke osjetljivosti uz istovremeno maksimiziranje specifičnosti na svakoj željenoj razini osjetljivosti, što omogućava fleksibilniju kliničku primjenu. Srednja vrijednost i standardna pogreška ciljanih metrika dobivene su prosjekom rezultata iz 50 vanjskih skupova testnih podataka (vanjska unakrsna validacija s pet odjeljaka ponovljena 10 puta). Najbolji model odabran je na temelju najviše postignute AUC vrijednosti.

Modeliranje je provedeno korištenjem programskih jezika Python (verzija 3.9.5, Python Software Foundation, Wilmington, DE, SAD) i R (verzija v2025.05.0 R Core Team, 2023, Beč, Austrija). Korišteni Python paketi bili su “numpy”, “pandas”, “scikit-learn” i “xgboost”. Unutar R ekosustava korištene su biblioteke „caret“, „glmnet“, „fastshap“, „pROC“, „ROCR“, „reshape2“, “dplyr”, “tidyr”, “purrr”, “ggplot2”, “plotly”, “tidymodels”, “ranger”, “xgboost” i “fastshap”. Postavljen je globalni *seed* za reproducibilnost rezultata u slučajevima u kojima funkcije koriste nasumičnost.

3.3. Statistička analiza podataka

Normalnost distribucije podataka testirana je Kolmogorov–Smirnovljevim testom. Za usporedbu dviju grupa, značajnost razlika procijenjena je t-testom za podatke s normalnom distribucijom ili Mann–Whitneyjevim testom za podatke koji odstupaju od normalnosti. Za usporedbe koje uključuju više od dvije grupe, korištena je ANOVA kada su podaci iz svih grupa bili normalno distribuirani, a Kruskal–Wallisov test u suprotnom slučaju. Za kategorijske značajke korišten je hi-kvadrat test. Statistička značajnost postavljena je na vrijednost $p < 0,05$. Za sve statističke analize i vizualizacije korišten je programski jezik R.

Provedena su dva pristupa bootstrap validacije za procjenu robusnosti modela. Klasična bootstrap analiza kojom je procijenjena varijabilnost performansi pri ponovnom učenju na različitim podskupovima iste populacije, što procjenjuje stabilnost modela u odnosu na fluktuacije uzorkovanja. Dodatno je napravljena klaster *bootstrap* analiza vanjskih preklopa iz ugniježdene unakrsne validacije što reflektira varijabilnost koja proizlazi iz različitih mogućih podjela podataka na trening i test setove.

3.4. Objašnjivost modela

SHAP (SHapley Additive exPlanations) metoda korištena je za dubinsku analizu važnosti kliničkih i laboratorijskih prediktora te za povećanje objašnjivosti i transparentnosti razvijenih modela za predikciju apendicitisa. Analiza je provedena na modelima koji su postigli optimalne performanse tijekom ugniježdene unakrsne validacije, pri čemu su SHAP vrijednosti izračunate kako za testne skupove unutar svakog vanjskog preklopa, tako i za konačni model treniran na potpunom skupu podataka. Kroz kvantifikaciju Shapley vrijednosti, detaljno je analizirano koliko svaka klinička varijabla (uključujući dob, trajanje simptoma, laboratorijske parametre poput leukocita, CRP-a, neutrofila i limfocita, te klinička ispitivanja) pridonosi konačnoj odluci modela za klasifikaciju pacijenta kao pozitivnog ili negativnog na apendicitis. Nadalje, SHAP analiza omogućila je vizualizaciju smjera utjecaja pojedinih varijabli (pozitivan ili negativan doprinos predikciji apendicitisa), identifikaciju međusobnih interakcija između kliničkih prediktora, te kvantifikaciju ukupne važnosti svake varijable kroz agregaciju individualnih Shapley vrijednosti preko svih predikcija. Ova metodologija posebno je vrijedna u medicinskom kontekstu jer omogućava kliničarima razumijevanje koje specifične kombinacije kliničkih nalaza i laboratorijskih vrijednosti model koristi za donošenje dijagnostičkih odluka, što povećava povjerenje u model i olakšava njegovu implementaciju u kliničku praksu. Svi izračuni provedeni su u programskom okruženju R korištenjem biblioteke “fastshap” optimizirane za rad s različitim tipovima strojnog učenja modela korištenih u ovom istraživanju.

4. REZULTATI

4.1. Rezultati istraživanja u pedijatrijskoj populaciji

Ukupno je uključeno 551 dijete operirano pod kliničkom slikom akutnog apendicitisa u izradi modela koji su bili stratificirani u skupine ovisno o PHD nalazu. Uključena su 22 parametra u izradu modela koje su uključivale laboratorijske vrijednosti, anamnestičke i kliničke podatke. Od ukupnog broja ispitanika, 47 je imalo negativan PHD nalaz, dok je u 252 slučaja opisan je nekomplikirani oblik apendicitisa i u 252 slučaja opisan je komplicirani oblik apendicitisa. Ovakav nesrazmjer među ispitivanim skupinama ukazuje na nebalansirani skup podataka. Karakteristike bolesnika za razvoj modela prikazane su u **Tablici 5**.

Za neke od navedenih varijabli nije bilo moguće u potpunosti pronaći vrijednosti, među kojima se ističu TRL, postotak limfocita, NRL i natrij kao vrijednosti s višim postotkom nedostajućih vrijednosti, dok se CRP, leukociti, spol, dob ističu kao vrijednosti bez nedostajućih vrijednosti (**Tablica 6**).

Tablica 5. Prikaz karakteristika ispitanika u pedijatrijskoj populaciji nakon stratifikacije po PHD nalazu

Parametar	Inocentni apendiks (n=47)	Nekomplicirani (n=252)	Komplicirani (n=252)	P
Dob, godine	11,63 ± 3,75	11,73 ± 3,63	11,75 ± 3,92	0,98*
Spol, n (%)				
Muško	27 (57,4%)	85 (33,7%)	84 (33,3%)	0,005*
Žensko	20 (42,6%)	167 (66,3%)	168 (66,7%)	
Visina, cm	153,32 ± 21,58	153,49 ± 20,34	155,96 ± 21,39	0,42*
Težina, kg	48,76 ± 19,58	46,57 ± 18,25	47,17 ± 20,12	0,46*
BMI, kg/m ²	18,76 ± 4,47	18,91 ± 3,8	19,06 ± 3,8	0,87*
Trajanje simptoma, sati	28 (31)	24 (15)	30 (24)	<0,001*
Tjelesna temperatura, °C	36,9 (1,2)	37 (0,9)	37,5 (1,4)	<0,001*
Mučnina, n (%)				
Da	30 (63,8%)	161 (63,9%)	198 (78,6%)	<0,001†
Ne	15 (31,9%)	83 (32,9%)	54 (21,4%)	
Nedostajući	2 (4,3%)	8 (3,2%)	0 (0%)	
Povraćanje, n (%)				
Da	16 (34%)	108 (42,9%)	172 (68,2%)	<0,001†
Ne	30 (63,8%)	142 (56,3%)	75 (29,8%)	
Nedostajući	1 (2%)	2 (0,8%)	5 (2%)	
Povratna osjetljivost, n (%)				
Nema	6 (12,8%)	29 (11,5%)	20 (7,9%)	<0,001†
Blago	14 (29,8%)	67 (26,6%)	36 (14,3%)	
Srednje	15 (31,9%)	110 (43,7%)	100 (39,7%)	
Jako	11 (23,4%)	42 (16,7%)	96 (38,1%)	
Nedostajući	1 (2,1%)	4 (1,5%)	0 (0%)	
Migracija boli, n (%)				
Da	19 (40,4%)	140 (55,6%)	140 (55,6%)	0,1†
Ne	27 (57,5%)	109 (43,3%)	99 (39,3%)	
Nedostajući	1 (2,1%)	3 (1,1%)	13 (5,1%)	
Leukociti, 10 ⁹ /L	11,5 ± 4,61	13,36 ± 4,36	16,58 ± 5,07	<0,001*
Neutrofili, %	76,9 (10,55)	79,3 (12,2)	84,9 (6,62)	<0,001*
Limfociti, %	15,6 (14,65)	13,8 (11,18)	7,95 (5,83)	<0,001*
Trombociti, 10 ⁹ /L	276,15 ± 71,86	274,66 ± 67,76	289,93 ± 72,07	0,05*
NLR	5,06 (6,22)	5,66 (5,3)	10,64 (8,66)	<0,001*
TLR	1,58 (0,64)	1,5 (0,96)	2,19 (1,42)	<0,001*
RDW, %	12,7 (0,9)	13 (1)	12,8 (1)	0,07*
MCHC, g/L	343,5 (11,25)	343 (14)	345 (11,25)	0,15*
MPV, fL	8,5 (2,9)	8,1 (2,13)	8,3 (2,05)	0,86*
CRP, mg/L	10,3 (47,35)	11,55 (23,3)	46,45 (66,25)	<0,001*
Natrij, mmol/L	140 (2)	139 (3)	137 (4)	<0,001*

Numerički podaci su prikazani kao srednja vrijednost ± standardna devijacija ili medijan (IQR), dok su kategorijski prikazani kao cijeli broj (postotak). Kratice: BMI: indeks tjelesne mase (engl. *body mass index*); MCHC: prosječna koncentracija hemoglobina u eritrocitima (engl. *mean corpuscular hemoglobin concentration*); MPV: prosječni volumen eritrocita (engl. *mean platelet volume*); NRL: omjer neutrofila i limfocita (engl. *neutrophil to lymphocyte ratio*); TLR: omjer trombocita i limfocita (engl. *thrombocyte to lymphocyte ratio*); RDW: raspodjela eritrocita po volumenu (engl. *red blood cell distribution width*)

*Jednosmjerna ANOVA ili Kruskal-Wallis test † Hi kvadrat test

Tablica 6. Prikaz nedostajućih vrijednosti u bazi podataka u pedijatrijskoj populaciji

Vrijednost	Postotak nedostajućih vrijednosti
TLR	29,2%
Postotak limfocita	29%
NLR	29%
Koncentracija natrija	24,5%
Visina	11,4%
Težina	11,4%
BMI	11,4%
RDW	8,5%
MPV	8,5%
MCHC	7,6%
Trombociti	5,3%
Postotak neutrofila	5,1%
Tjelesna temperatura	3,4%
Mučnina	3,4%
Migracija boli	3,1%
Povraćanje	1,5%
Duljina trajanja simptoma	1,1%
Rebound tenderness	0,9%
Dob	0%
Spol	0%
Leukociti	0%
CRP	0%

Kratice: BMI: indeks tjelesne mase (engl. *body mass index*); MCHC: prosječna koncentracija hemoglobina u eritrocitima (engl. *mean corpuscular hemoglobin concentration*); MPV: prosječni volumen eritrocita (engl. *mean platelet volume*); NLR: omjer neutrofila i limfocita (engl. *neutrophil to lymphocyte ratio*); RDW: raspodjela eritrocita po volumenu (engl. *red blood cell distribution width*); TLR: omjer trombocita i limfocita (engl. *thrombocyte to lymphocyte ratio*)

Model za predikciju apendicitisa

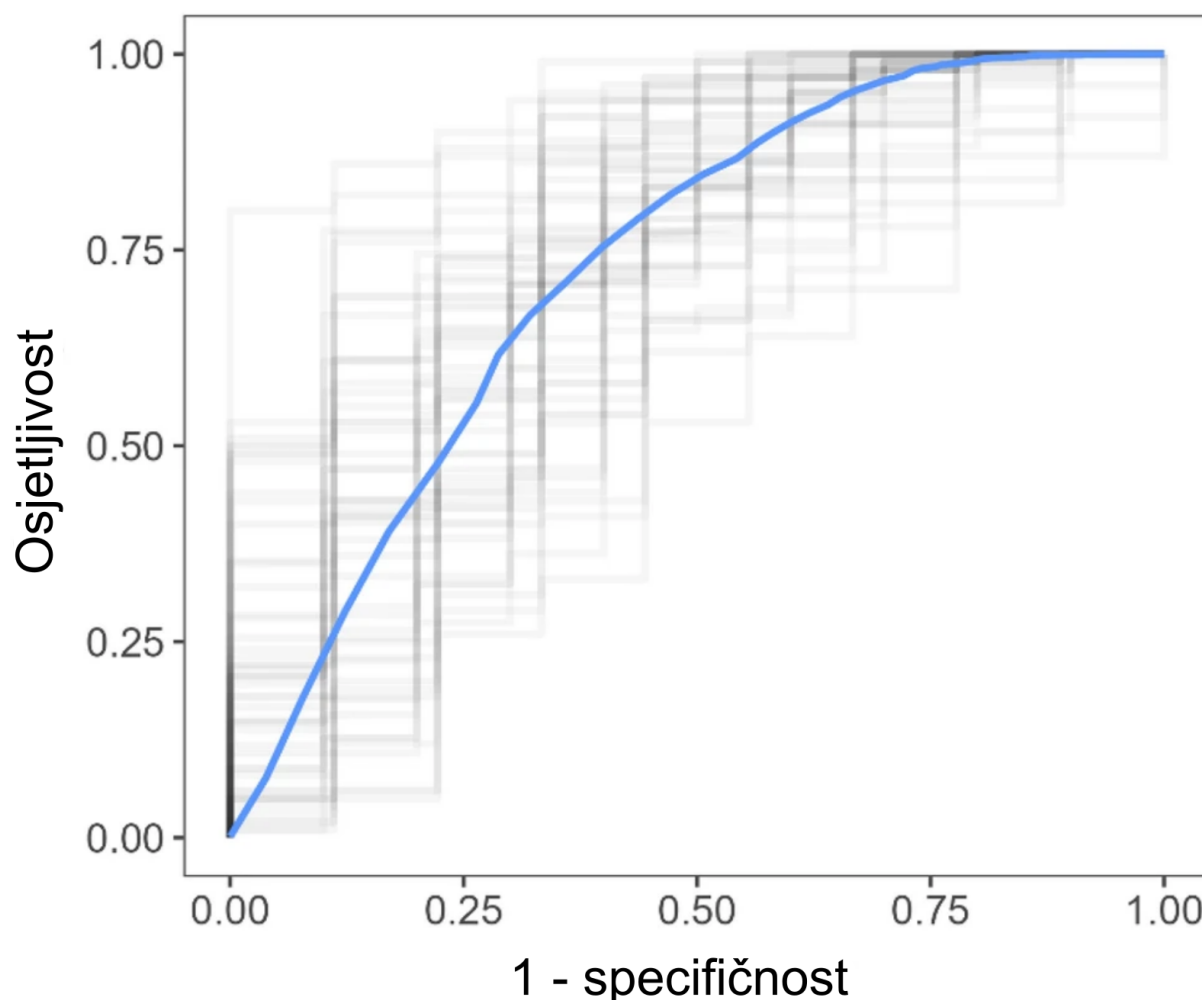
U svrhu predikcije apendicitisa, negativni nalazi PHD-a klasificirani su kao nule, a pozitivni (nekomplikirani i komplicirani) kao jedinice, nakon čega su razvijeni modeli temeljeni na ovoj binarnoj klasifikaciji. Najbolji rezultati postignuti su modelom slučajnih šuma s prosječnom osjetljivošću od $0,997 \pm 0,001$ i specifičnošću od $0,17 \pm 0,01$ (**tablica 7.**, **slika 13.**).

Tablica 7. Performanse modela u predikciji akutnog apendicitisa

Model	Osjetljivost	Specifičnost
Slučajne šume	$0,997 \pm 0,001$	$0,17 \pm 0,01$
XGBoost	$0,9982 \pm 0,0006$	$0,12 \pm 0,01$
Logistička regresija	$0,997 \pm 0,001$	$0,052 \pm 0,006$

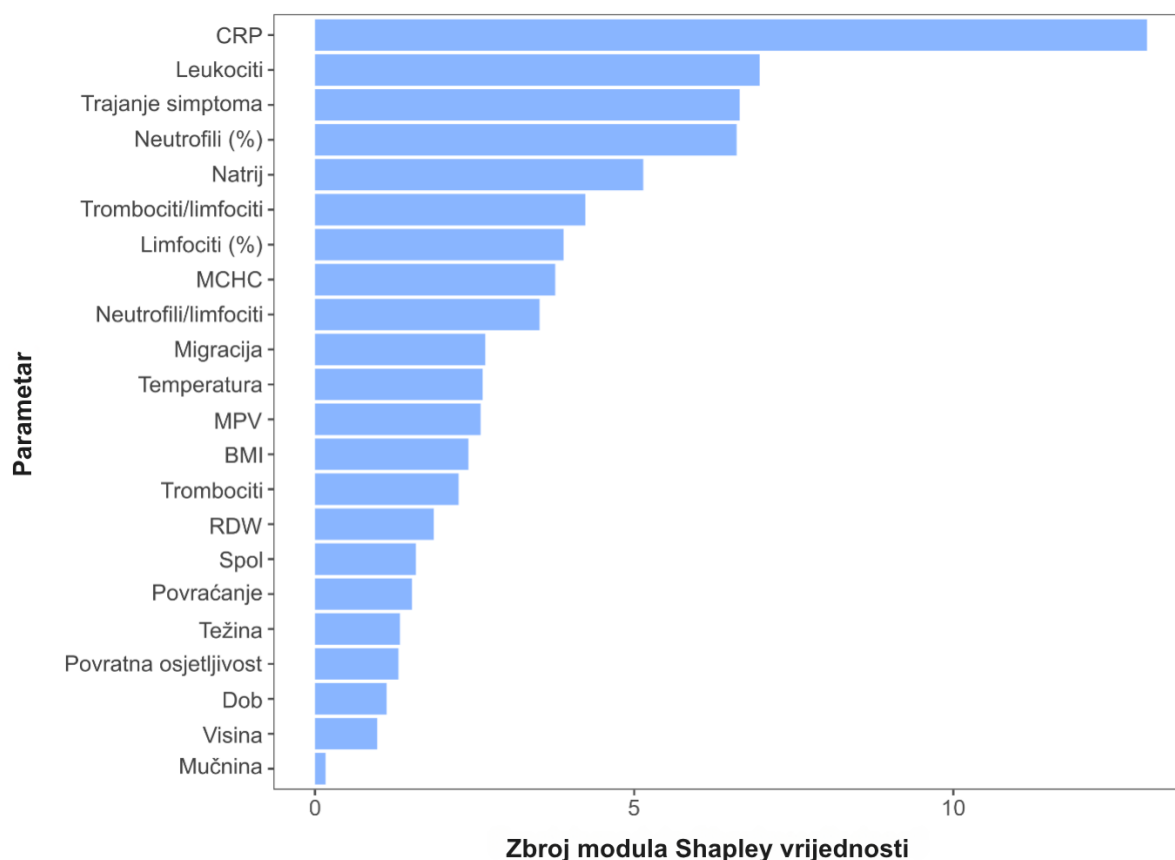
Podaci su prikazani kao prosjek $\pm$ standardna pogreška

Od lažno negativnih rezultata, gotovo svi su bili slučajevi nekomplikiranog apendicitisa, a ne komplicirani oblici bolesti. Ovu razliku možemo kvantificirati zasebnim izračunavanjem osjetljivosti za ove dvije skupine bolesnika. Osjetljivost za detekciju nekomplikiranog oblika bolesti iznosila je $0,995 \pm 0,002$, dok je za detekciju kompliciranog oblika bolesti bila još viša, na razini od $0,9996 \pm 0,0004$.



Slika 13. ROC krivulja za prosječnu vrijednost modela slučajnih šuma u predikciji akutnog apendicitisa (plava krivulja) i ROC krivulje za svaki vanjski preklop (sive krivulje).

Kako bismo bolje razumjeli kako je model obradio informacije, izračunate su Shapleyjeve vrijednosti na jednom od vanjskih preklopa kako bismo odredili važnost pojedinih značajki (**Slika 14.**). Među najznačajnijima se izdvajaju CRP, leukociti, duljina trajanja simptoma, postotak neutrofila i natrij.



Slika 14. Prikaz važnosti pojedinih značajki modela slučajnih šuma u predikciji apendicitisa prema Shapleyjevim vrijednostima. Prikazano kao zbroj modula Shapley vrijednosti za svaki vanjski preklop. Kratice: CRP: C-reaktivni protein; MCHC: prosječna koncentracija hemoglobina u eritrocitima (engl. *mean corpuscular hemoglobin concentration*); MPV: prosječni volumen eritrocita (engl. *Mean platelet volume*); BMI: indeks tjelesne mase (engl. *Body mass indeks*); RDW: raspodjela eritrocita po volumenu (engl. *red blood cell distribution width*)

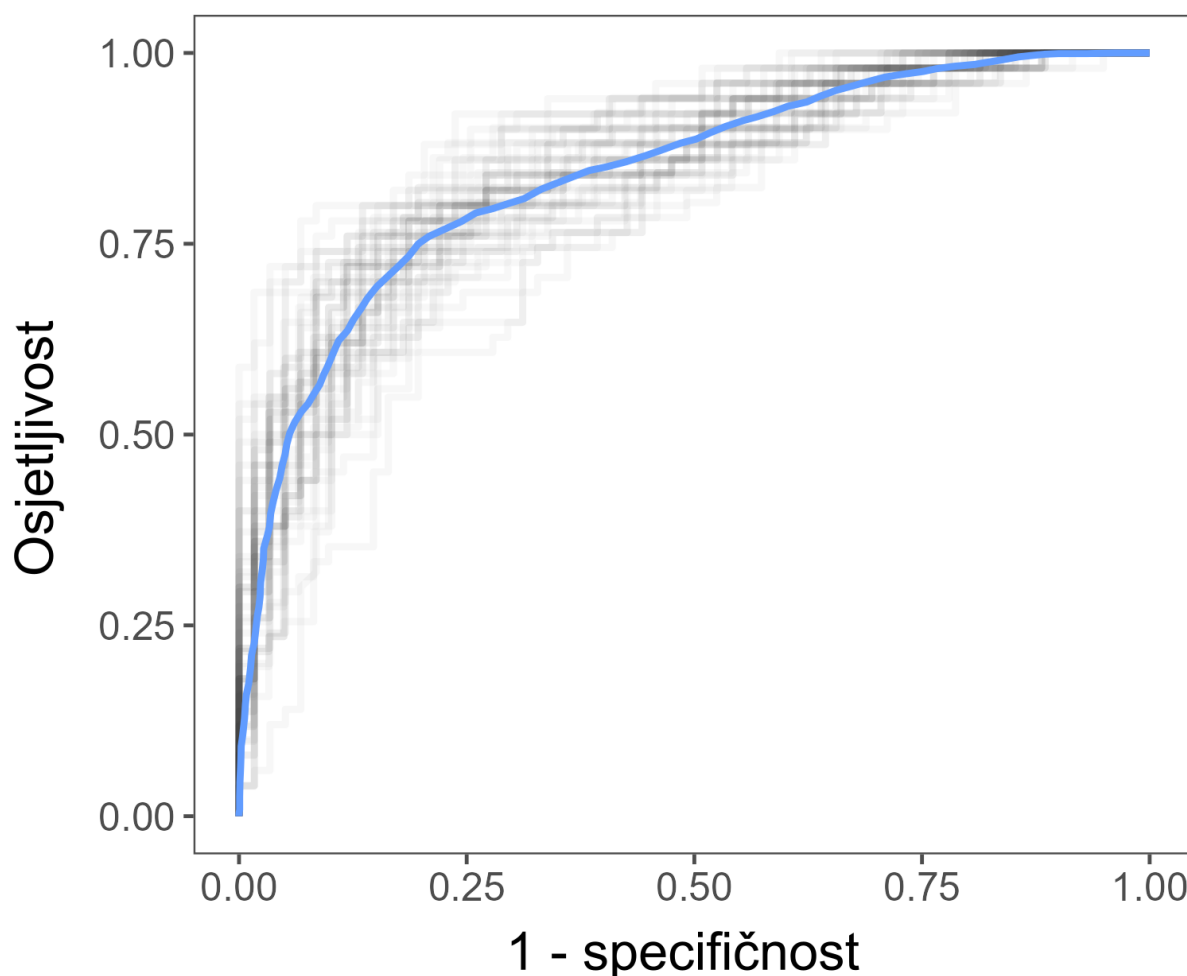
Model za predikciju kompliciranog apendicitisa

Uzimajući u obzir mogućnost konzervativnog antibiotskog liječenja, razvijen je model za predikciju kompliciranih oblika apendicitisa. Zadržavajući sve ostale parametre identičnima kao u prethodno opisanom pristupu detekcije akutnog apendicitisa, spojeni su negativni PHD nalazi i PHD nalazi nekompliciranog apendicitisa u jednu kategoriju, dok se druga sastojala od kompliciranih oblika bolesti. Model slučajnih šuma ponovno se pokazao najboljim sa srednjom osjetljivošću od $0,994 \pm 0,002$ i specifičnošću od $0,129 \pm 0,009$ (**Tablica 8, Slika 15**).

Tablica 8. Performanse modela u predikciji kompliciranog oblika apendicitisa

Model	Osjetljivost	Specifičnost
Slučajne šume	0,994 ± 0,002	0,129 ± 0,009
XGBoost	0,963 ± 0,008	0,26 ± 0,02
Logistička regresija	0,998 ± 0,0007	0,009 ± 0,004

Podaci su prikazani kao prosjek ± standardna pogreška

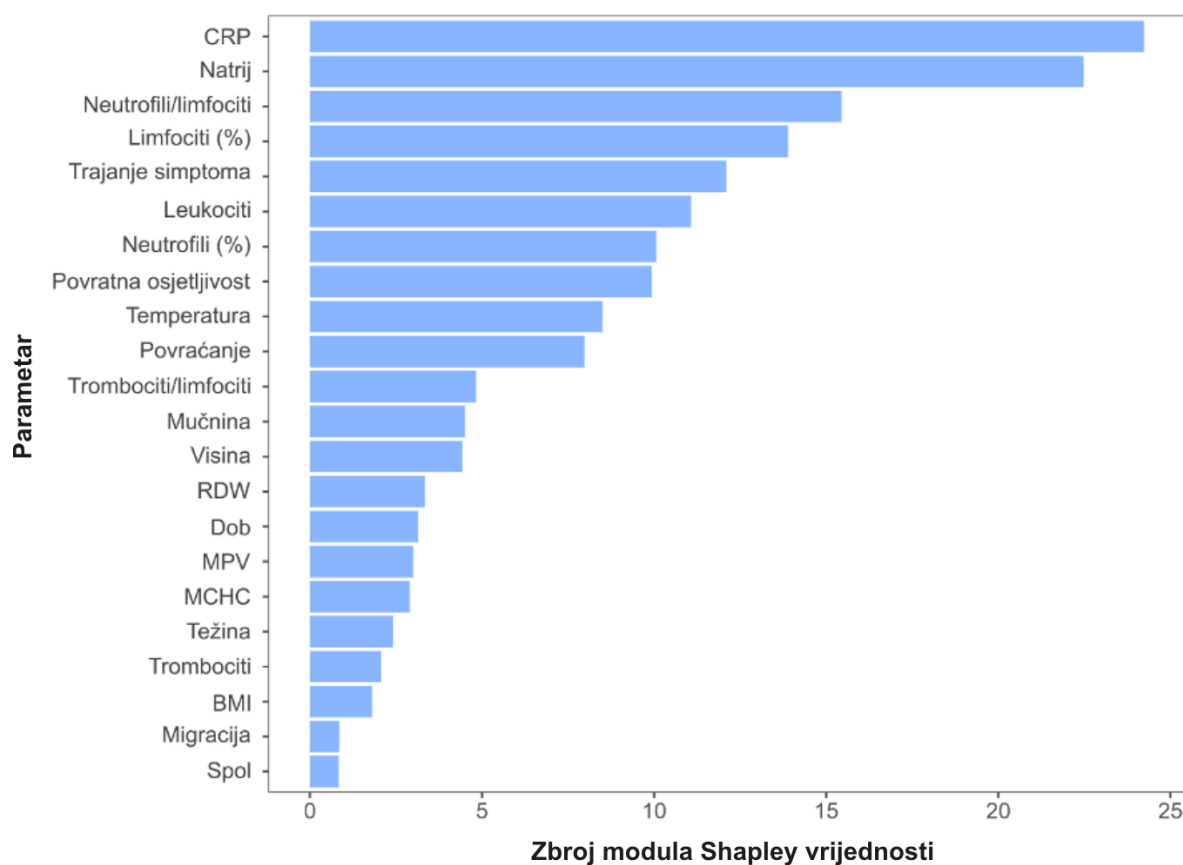


Slika 15. ROC krivulja za prosječnu vrijednost modela slučajnih šuma u predikciji kompliciranog oblika apendicitisa (plava krivulja) i ROC krivulje za svaki vanjski preklop (sive krivulje).

Kada se u ovom modelu razdvoje specifičnosti za negativne i nekomplikirane oblike bolesti, dobije se specifičnost od $0,25 \pm 0,02$ za negativne PHD nalaze i $0,107 \pm 0,008$ za nekomplikirane oblike bolesti.

Nadalje, ponovno su izračunate Shapleyjeve vrijednosti kako bi se odredila važnost određenih značajki u modelu (**Slika 16.**). Među navedenim značajkama najistaknutije su CRP,

natrij, omjer neutrofila i limfocita, postotak limfocita, duljina trajanja simptoma, leukociti, postotak neutrofila.



Slika 16. Prikaz važnosti pojedinih značajki modela slučajnih šuma u predikciji kompliciranog apendicitisa prema Shapleyjevim vrijednostima. Prikazano kao zbroj modula Shapley vrijednosti od svakog vanjskog preklopa. Kratice: CRP: C-reaktivni protein; MCHC: prosječna koncentracija hemoglobina u eritrocitima (engl. *mean corpuscular hemoglobin concentration*); MPV: prosječni volumen eritrocita (engl. *Mean platelet volume*); BMI: indeks tjelesne mase (engl. *Body mass indeks*); RDW: raspodjela eritrocita po volumenu (engl. *red blood cell distribution width*)

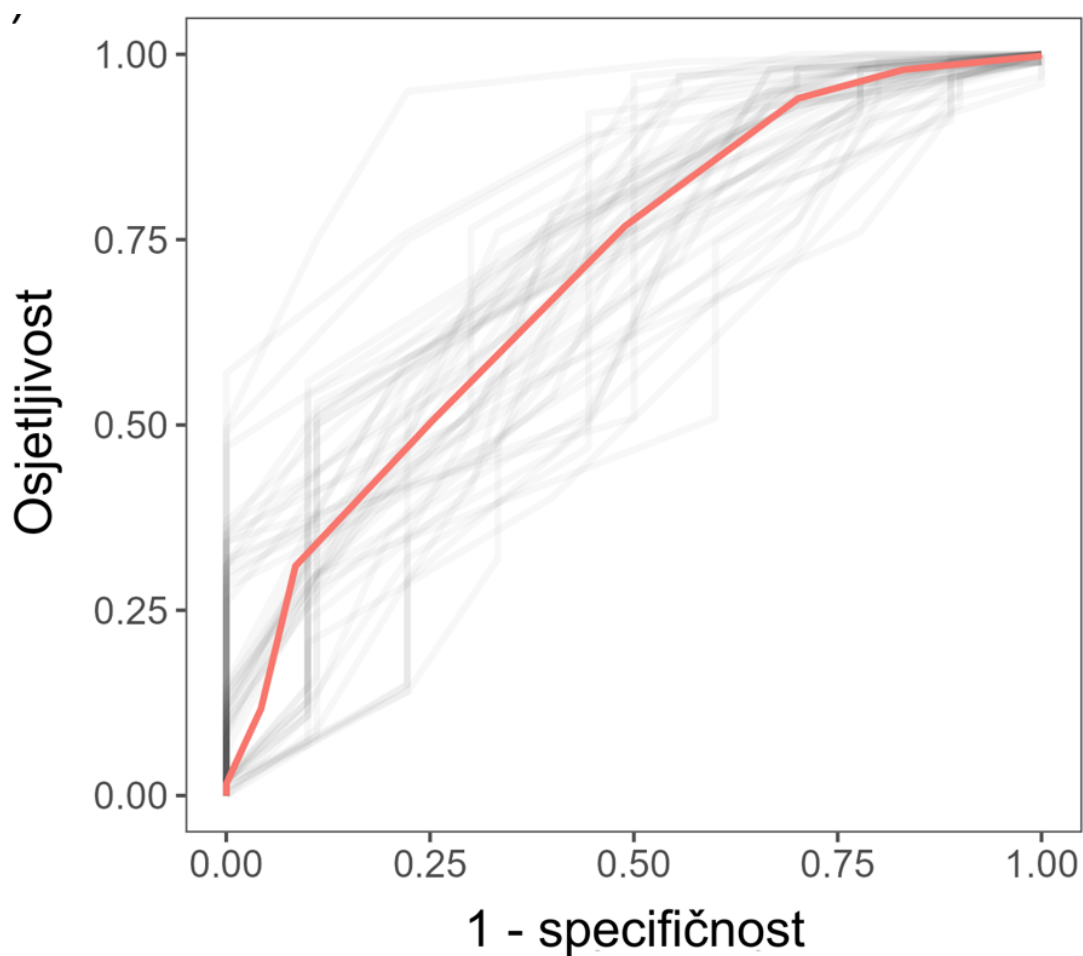
Usporedba s *Appendicitis inflammatory response* (AIR) sustavom bodovanja

Kako bi se ustanovile prednosti ovog modela, izračunate su vrijednosti osjetljivosti i specifičnosti za AIR za sve moguće bodovne pragove, pri čemu su se koristile prosječne vrijednosti kroz svih istih 50 vanjskih preklopa korištenih u evaluaciji razvijenog modela (Tablica 9., slika 17.).

Tablica 9. Prikaz osjetljivosti i specifičnosti AIR bodovne tablice za svaki bodovni prag

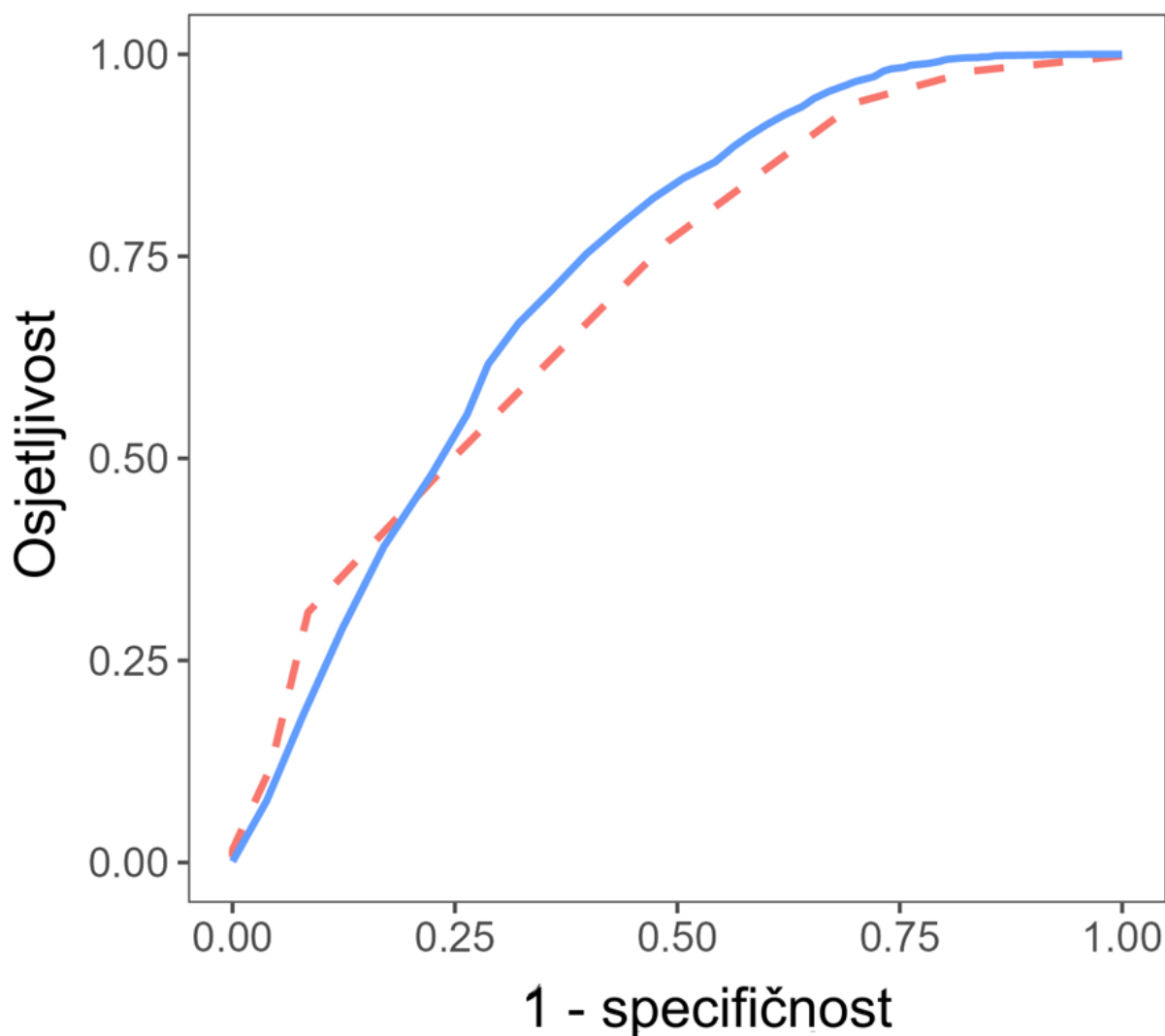
Bodovni prag	Osjetljivost	Specifičnost
0	$0,9980 \pm 0,0006$	0 ± 0
1	$0,979 \pm 0,002$	$0,17 \pm 0,01$
2	$0,940 \pm 0,003$	$0,30 \pm 0,02$
3	$0,768 \pm 0,004$	$0,51 \pm 0,02$
4	$0,507 \pm 0,006$	$0,75 \pm 0,02$
5	$0,310 \pm 0,005$	$0,92 \pm 0,02$
6	$0,117 \pm 0,004$	$0,96 \pm 0,01$
7	$0,016 \pm 0,002$	1 ± 0
8	0 ± 0	1 ± 0
9	0 ± 0	1 ± 0
10	0 ± 0	1 ± 0
11	0 ± 0	1 ± 0
12	0 ± 0	1 ± 0

Podaci su prikazani kao prosjek $\pm$ standardna pogreška



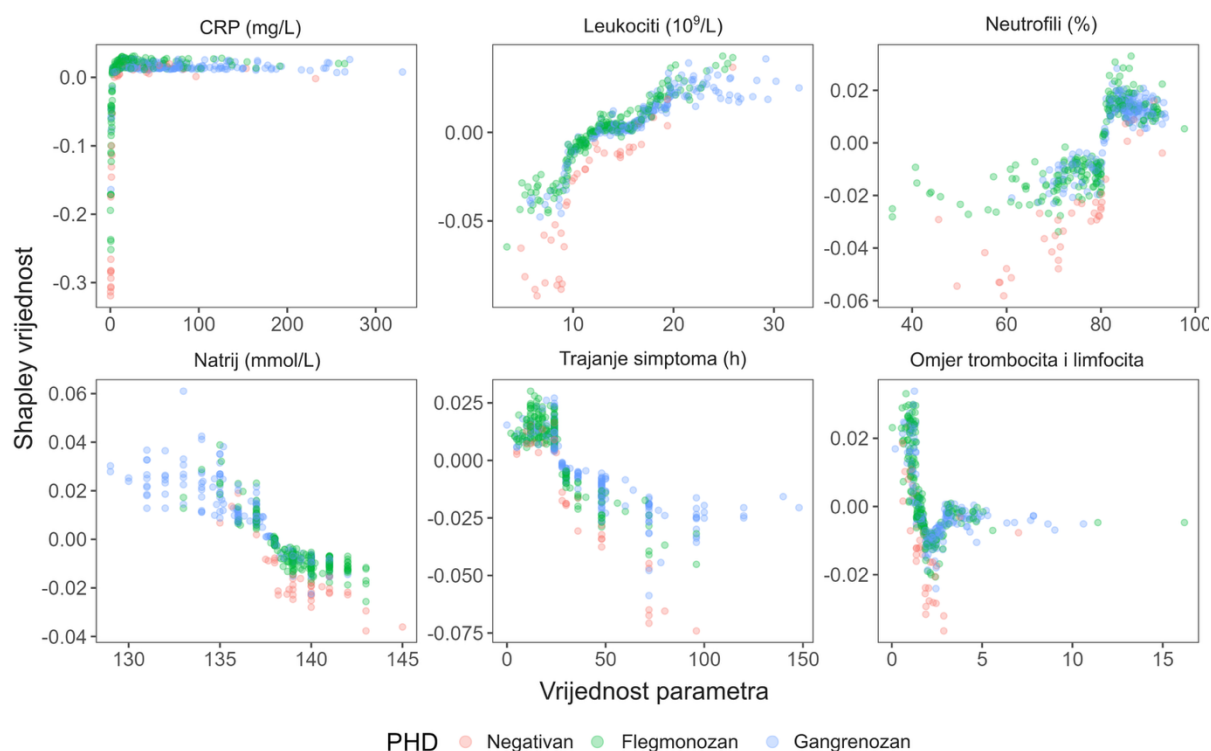
Slika 17. ROC krivulja za prosječnu vrijednost AIR bodovnog sustava (crvena) i ROC vrijednosti za svaki vanjski preklop (sivo)

U pogledu osjetljivosti, učinkovitost pri pragu AIR sustava bodovanja većem od 0 bila je slična našoj, no specifičnost je iznosila 0. Za prag AIR-a postavljen na vrijednosti veće od 1, specifičnost je bila usporediva s onom našeg modela, ali je osjetljivost bila značajno niža ($p < 0,001$, t-test). Izvan područja visoke osjetljivosti, naš je model nadmašio AIR i u većini drugih kombinacija osjetljivosti–specifičnosti (**Slika 18.**).



Slika 18. Usporedba ROC krivulja modela slučajnih šuma (plava krivulja) i AIR sustava bodovanja (crvena isprekidana krivulja).

U pogledu pragova pri kojima vrijednosti parametara povećavaju vjerojatnost akutnog apendicitisa, primjećujemo sličnost s graničnim vrijednostima AIR sustava bodovanja: oko 10 mg/L za CRP, $10 \times 10^9/L$ za broj leukocita te 80% za postotak neutrofila (**Slika 19.**).



Slika 19. Odnos najvažnijih prediktora i PHD nalaza prema Shapley vrijednostima.

4.2. Rezultati istraživanja u odrasloj populaciji

Provedena je sveobuhvatna analiza nedostajućih podataka. Postotak nedostajućih vrijednosti kretao se u rasponu od 0% do 5,7% kroz sve varijable. Razine natrija pokazale su najveći udio nedostajućih vrijednosti (5,7%), nakon čega slijedi tjelesna temperatura (4,4%), a još se izdvajaju mučnina, povraćanje, trajanje simptoma, povratna osjetljivost, migracija boli i CRP. Sve ostale varijable nisu imale nedostajućih vrijednosti (**Tablica 10**).

Komparabilno s rezultatima u pedijatrijskog populaciji, u svrhu predikcije apendicitisa, negativni nalazi PHD-a (bez znakova akutnog apendicitisa) klasificirani su kao nule, a pozitivni (nekomplikirani i komplikirani) kao jedinice, nakon čega su razvijeni modeli temeljeni na ovoj binarnoj klasifikaciji. Karakteristike bolesnika po ovakvoj klasifikaciji su prikazane u **Tablici 11**.

Tablica 10. Prikaz nedostajućih vrijednosti u bazi podataka u odrasloj populaciji

Vrijednost	Postotak nedostajućih vrijednosti
Koncentracija natrija	5,7%
Tjelesna temperatura	4,4%
Mučnina	2,8%
Povraćanje	1,7%
Trajanje simptoma	1,7%
Rebound	0,9%
Migracija	0,7%
CRP	0,4%
Spol	0%
Bol u DDK	0%
Godine	0%
Leukociti	0%
MPV	0%
Postotak neutrofila	0%
Postotak limfocita	0%
TLR	0%
NLR	0%

Kratice: DDK: donji desni kvadrant; MPV: prosječni volumen eritrocita (engl. *mean platelet volume*); NRL: omjer neutrofila i limfocita (engl. *neutrophil to lymphocyte ratio*); TLR: omjer trombocita i limfocita (engl. *thrombocyte to lymphocyte ratio*)

Tri modela strojnog učenja, logistička regresija, slučajne šume i XGBoost, evaluirani su s posebnim naglaskom na osjetljivost za detekciju akutnog apendicitisa, zbog čega je rađeno pomicanje praga odluke kako bi se maksimizirala osjetljivost. Najbolji rezultati postignuti su modelom logističke s postignutim AUC-om od $0,765 \pm 0,001$. Napravljeni su višestruki modeli ovisno o minimalnoj traženoj osjetljivosti u kojima je prag odluke pomican kako bi se dobila osjetljivost od najmanje 0,98 do 0,995 (Tablica 12., slika 20.). Za najvišu dobivenu osjetljivost od 0,998 dobivena je specifičnost od 0,038.

Tablica 11. Prikaz karakteristika bolesnika u odrasloj populaciji nakon stratifikacije po PHD nalazu

Parametar	Inocentni apendiks (n=56)	Nekomplicirani (n=210)	Komplicirani (n= 276)	P
Dob, godine	32,5 (21)	34,5 (22)	45 (26,5)	<0,001*
Spol, n (%)				
Muško	20 (35,7%)	125 (59,5%)	152 (55,1%)	0,006†
Žensko	36 (64,3%)	85 (40,5%)	124 (44,9%)	
Trajanje simptoma, sati	32 (43)	20 (12)	36 (24)	<0,001*
Povišena tjelesna temperatura, n(%)				
Da	18 (32,1%)	37 (17,6%)	101 (36,6%)	<0,001†
Ne	36 (64,3%)	167 (79,5%)	159 (57,6%)	
Nedostajući	2 (3,6%)	6 (2,9%)	16 (5,8%)	
Mučnina, n (%)				
Da	26 (46,4%)	92 (43,8%)	144 (52,1%)	0,096†
Ne	30 (53,6%)	114 (54,3%)	121 (43,8%)	
Nedostajući	0 (0%)	4 (1,9%)	1 (0,1%)	
Povraćanje, n (%)				
Da	11 (19,6%)	58 (27,6%)	109 (39,5%)	0,001†
Ne	45 (80,4%)	150 (71,5%)	160 (57,9%)	
Nedostajući	0 (0%)	2 (0,9%)	7 (2,6%)	
Povratna osjetljivost, n (%)				
Da	45 (80,4%)	153 (72,9%)	225 (81,5%)	0,089†
Ne	11 (19,6%)	54 (25,7%)	49 (17,8%)	
Nedostajući	0 (0%)	3 (1,4%)	2 (0,7%)	
Migracija boli, n (%)				
Da	24 (42,9%)	104 (49,5%)	125 (45,3%)	0,623†
Ne	31 (55,4%)	106 (50,5%)	148 (53,6%)	
Nedostajući	1 (1,7%)	0 (0%)	3(1,1%)	
Bol u DDK, n(%)				
Da	53 (94,6%)	201 (95,7%)	253 (91,7%)	<0,001†
Ne	3 (5,4%)	1 (0,5%)	23 (8,3%)	
Nedostajući	0 (0%)	8 (3,8%)	0 (0%)	
Leukociti, 10 ⁹ /L	10,6 (5)	13,05 (5,4)	14,5 (6,1)	<0,001*
Neutrofili, %	76,05 (15,25)	80,5 (10,6)	83,9 (10,55)	<0,001*
Limfociti, %	14,9 (12,05)	12,2 (8,1)	9,5 (7,1)	<0,001*
MPV	9,4 (2,2)	9,55 (1,7)	9,45 (1,8)	0,462*
NLR	5,45 (4,43)	6,74 (5,69)	9,38 (8,97)	<0,001*
TLR	153,75 (111,46)	147,79 (92,01)	177,36 (124,55)	<0,001*
CRP, mg/L	36,8 (68,2)	22,1 (43,95)	84,3 (129,5)	<0,001*
Natrij, mmol/L	140 (3)	139 (3)	138 (4)	<0,001*

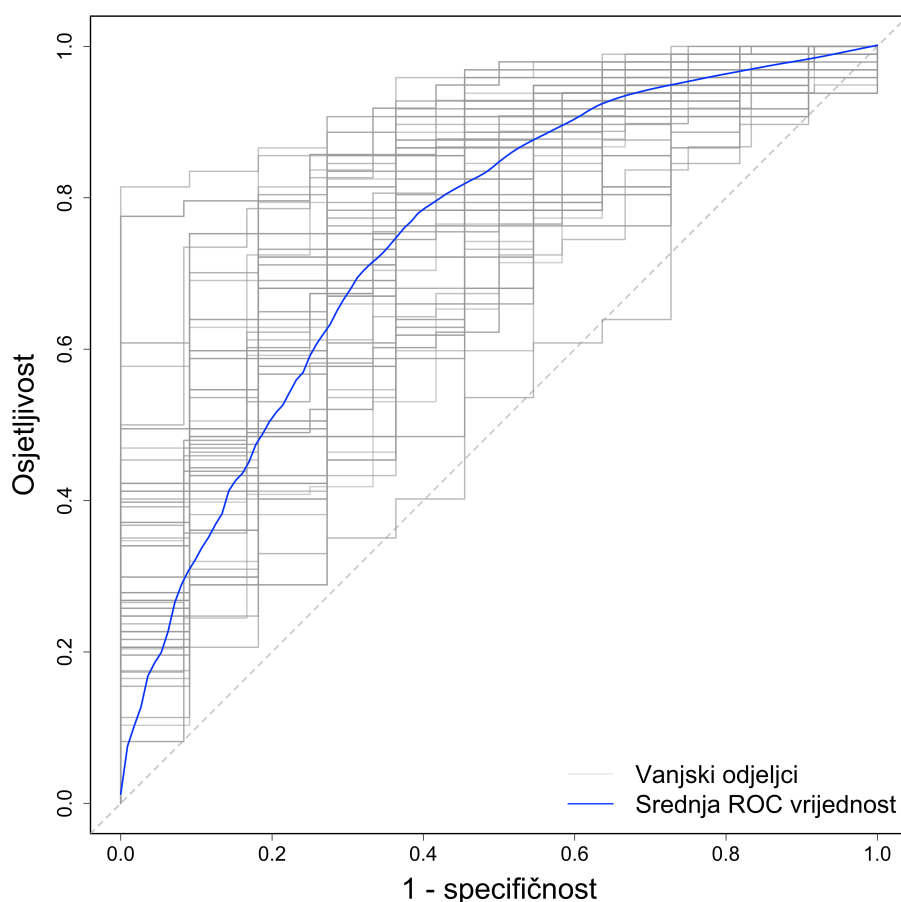
Numerički podaci su prikazani kao medijan (IQR), a kategorijski kao cijeli broj (postotak). Kratice: DDK: donji desni kvadrant; MPV: prosječni volumen eritrocita (engl. *mean platelet volume*); NLR: omjer neutrofila i limfocita (engl. *neutrophil to lymphocyte ratio*); TLR: omjer trombocita i limfocita (engl. *thrombocyte to lymphocyte ratio*)

* Kruskal-Wallis test † Hi kvadrat test

Tablica 12. Performanse modela u predikciji akutnog apendicitisa

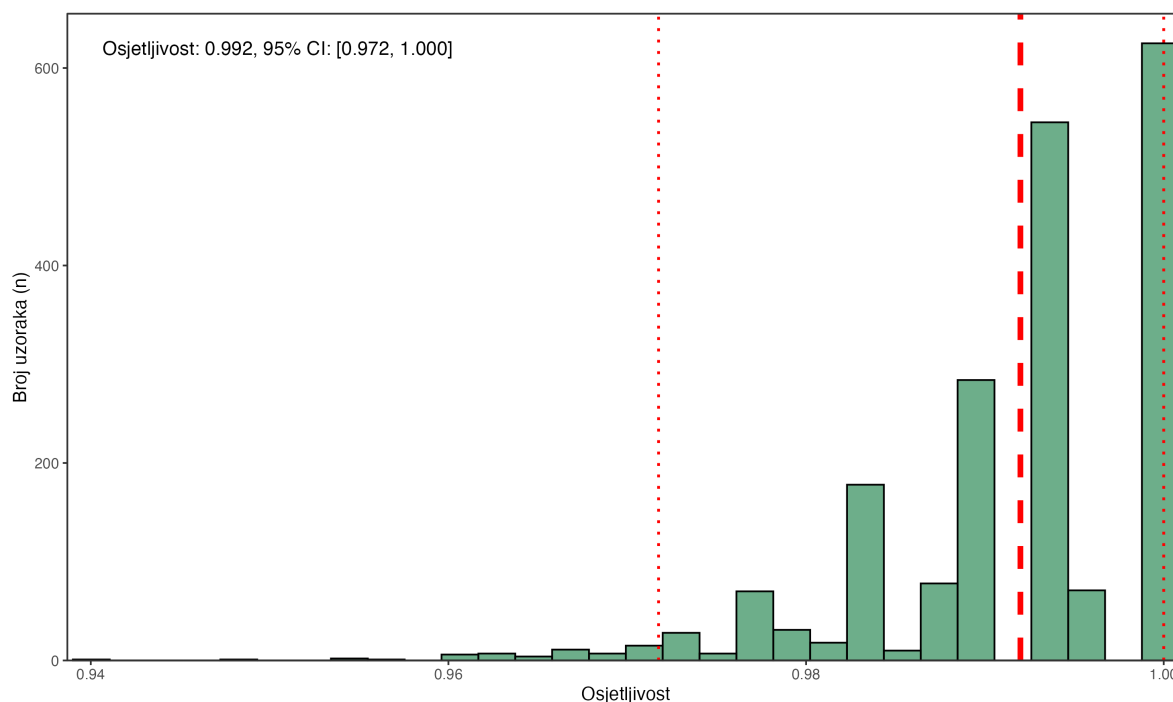
Model i prag odluke	AUC	Osjetljivost	Specifičnost
Slučajne šume			
PO1	0,745 ± 0,011	0,982 ± 0,002	0,100 ± 0,011
PO2		0,994 ± 0,001	0,049 ± 0,008
PO3		0,997 ± 0,001	0,022 ± 0,006
XGBoost			
PO1	0,723 ± 0,013	0,972 ± 0,004	0,146 ± 0,018
PO2		0,984 ± 0,004	0,094 ± 0,017
PO3		0,989 ± 0,004	0,062 ± 0,014
Logistička regresija			
PO1	0,765 ± 0,01	0,982 ± 0,002	0,102 ± 0,013
PO2		0,993 ± 0,002	0,07 ± 0,011
PO3		0,998 ± 0,001	0,038 ± 0,009

Podaci su prikazani kao prosjek ± standardna pogreška. Kratice: AUC – površina ispod ROC krivulje (engl. *Area under the curve*); PO1 – prag odluke pri kojem je osjetljivost minimalno 0,98; PO2 – prag odluke pri kojem je osjetljivost minimalno 0,99; PO3 – prag odluke pri kojem je osjetljivost >0,995

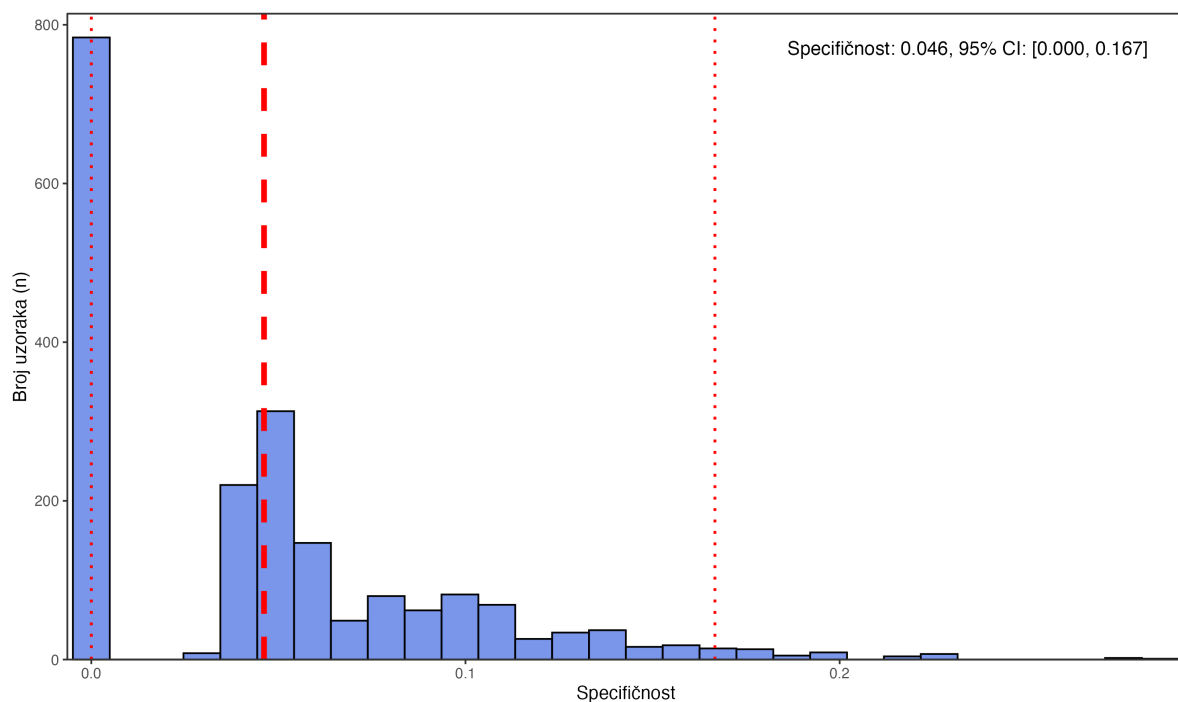


Slika 20. ROC krivulja za prosječnu vrijednost modela logističke regresije (plava krivulja) i ROC krivulje za svaki vanjski preklop (sive krivulje).

Kako bismo dodatno procijenili performanse logističke regresije za prag odluke u kojem je osjetljivost $>0,995$, provedene su dvije *bootstrap* analize. U prvoj je korišten klasični *bootstrap* u kojem je uzet originalni skup podataka i provedeno 2000 iteracija s *bootstrap* uzorkovanjem ispitanika na kojima je ponovno naučen već kreiran model te ga testira na *out-of-bag* ispitanicima tj. onima koji nisu ušli u *bootstrap* uzorak. To je napravljeno s ciljem procjene varijabilnosti performansi pri ponovnom učenju na različitim podskupovima iste populacije, što procjenjuje stabilnost modela u odnosu na fluktuacije uzorkovanja. Za odabrani prag odluke PO3, osjetljivost je bila 0,991 (95% CI 0,968-1), specifičnost 0,049 (95% CI 0-0,174), accuracy 0.894 (95% CI 0,86-0,928), PPV 0,9 (95% CI 0,866-0,935), NPV 0,424 (95% CI 0-1) i F1 ocjena 0,943 (95% CI 0,924-0,962) (Slika 21. i 22.). U drugoj analizi koristili smo već spremljene specifikacije modela stratificirano po vanjskim presjecima, a sam postupak je temeljen na klaster *bootstrappingu* vanjskih preklop dobivenih u sklopu već izvršene ugniježdene unakrsne validacije. Na taj način, intervali pouzdanosti dobiveni ovom metodom reflektiraju varijabilnost koja proizlazi iz različitih mogućih podjela podataka na trening i test skupove tj. varijabilnost u generalizaciji već odabranog modela i njegovog praga. Za odabrani prag odluke PO3, osjetljivost je bila 0,992 (95% CI 0,973-1), specifičnost 0,043 (95% CI 0-0,158), preciznost 0.894 (95% CI 0,861-0,928), PPV 0,9 (95% CI 0,866-0,935), NPV 0,409 (95% CI 0-1) i F1 ocjena 0,944 (95% CI 0,925-0,963).

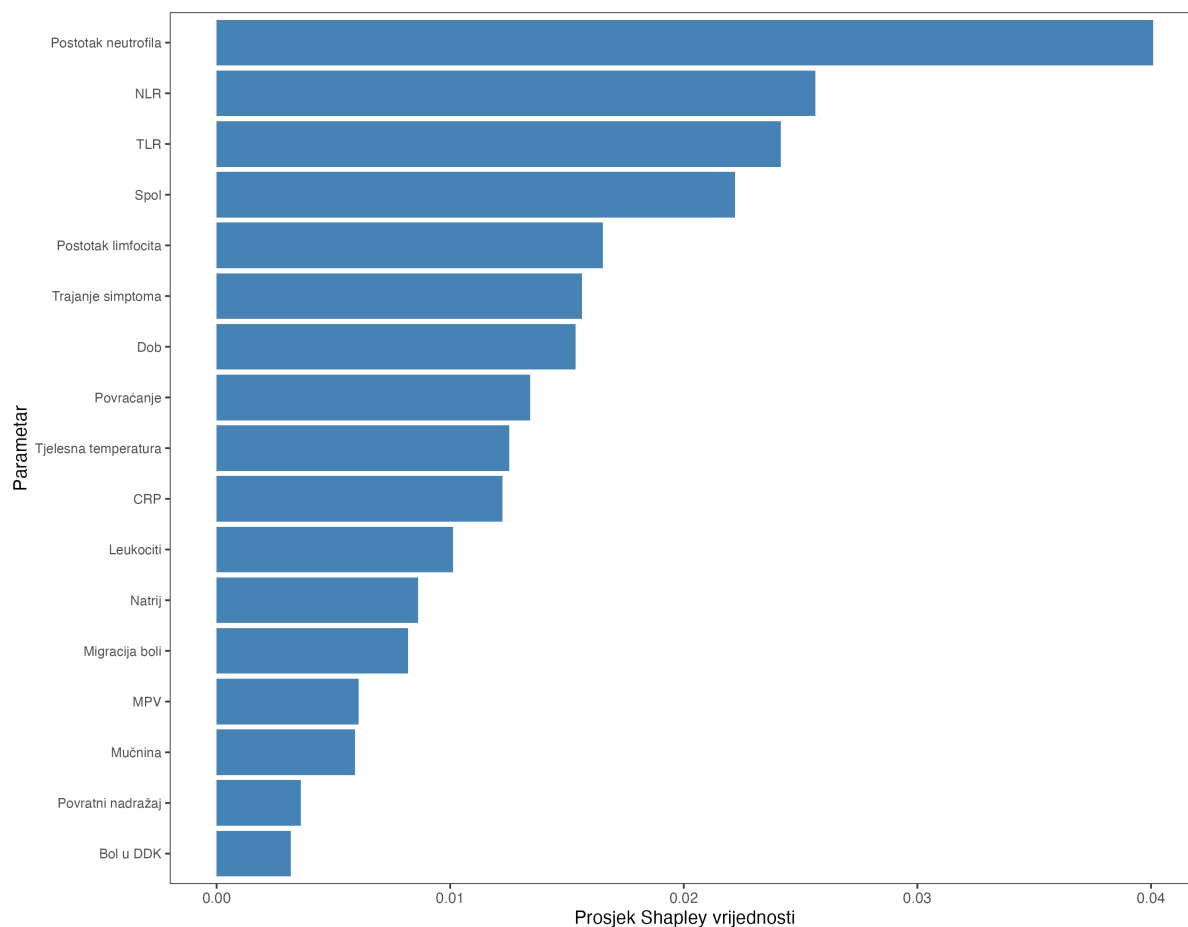


Slika 21. Prikaz prosječne osjetljivosti i 95% CI dobiven klasičnom *bootstrap* validacijom s 2000 iteracija za logističku regresiju na pragu odluke za koji je osjetljivost $>0,995$

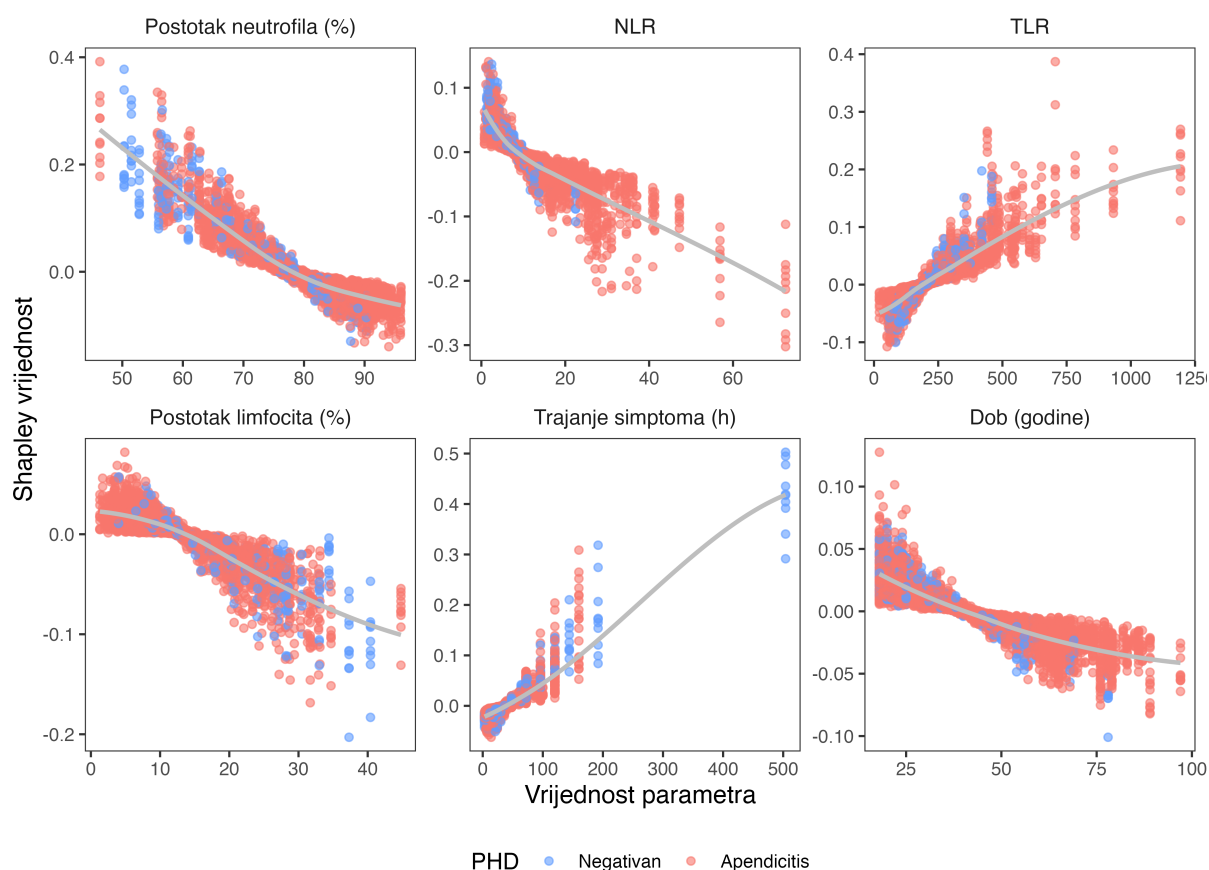


Slika 22. Prikaz prosječne specifičnosti i 95% CI dobiven klasičnom *bootstrap* validacijom s 2000 iteracija za logističku regresiju na pragu odluke za koji je osjetljivost $>0,995$

Kako bismo bolje razumjeli kako je model obradio informacije, izračunate su Shapleyjeve vrijednosti prema odabranom pragu odluke PO3 za koji je osjetljivost $>0,995$. Provedena je SHAP analiza s 10000 permutacija izračunatih na *out-of-sample* uzorcima na odabranom pragu odluke (**Slika 23.**). Također, na **Slici 24.** prikazan je odnos najvažnijih numeričkih prediktora i SHAP vrijednosti.



Slika 23. Prikaz važnosti pojedinih značajki logističke regresije prema Shapleyjevim vrijednostima za odabrani prag odluke za koji je osjetljivost $>0,995$. Prikazano kao prosjek modula Shapley vrijednosti za svaki preklop. Kratice: DDK: donji desni abdominalni kvadrant; MPV: prosječni volumen eritrocita (engl. *mean platelet volume*); NRL: omjer neutrofila i limfocita (engl. *neutrophil to lymphocyte ratio*); TLR: omjer trombocita i limfocita (engl. *thrombocyte to lymphocyte ratio*)



Slika 24. Odnos najvažnijih prediktora i PHD nalaza logističke regresije prema Shapley vrijednostima za prag odluke za koji je osjetljivost $>0,995$. Kratice: NLR: omjer neutrofila i limfocita (engl. *neutrophil to lymphocyte ratio*); TLR: omjer trombocita i limfocita (engl. *thrombocyte to lymphocyte ratio*)

Model za predikciju kompliciranog apendicitisa

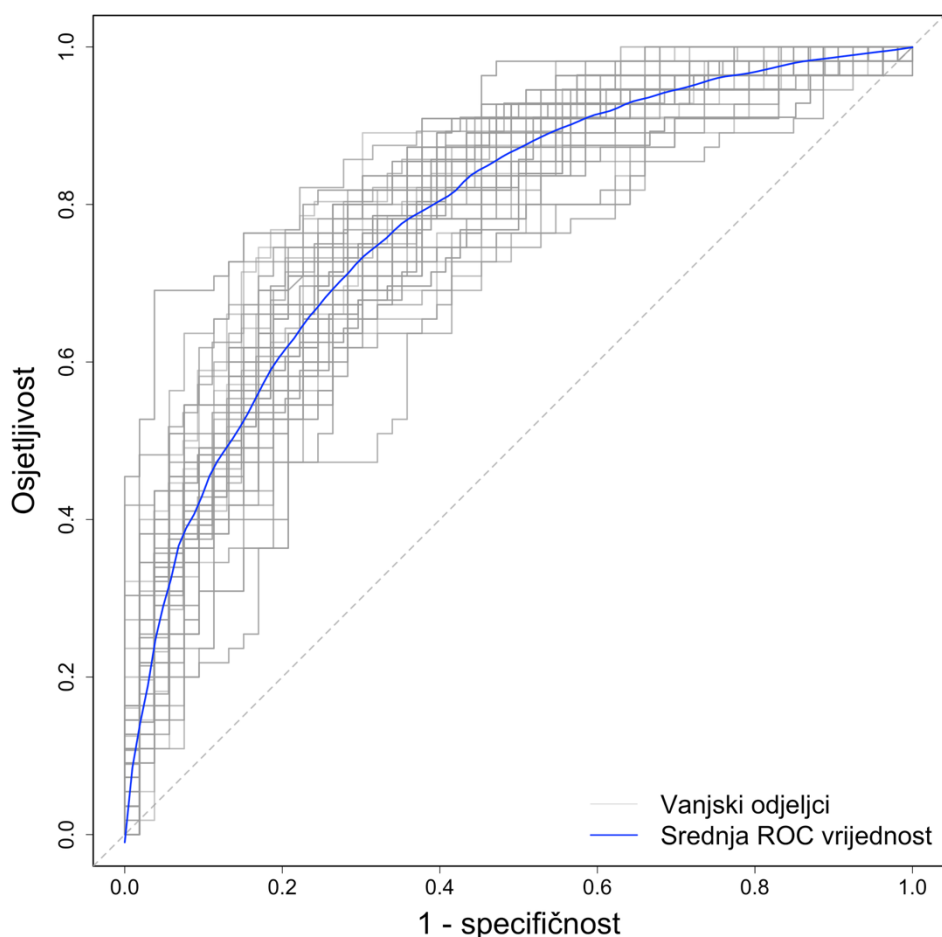
Komparabilno s modelima u pedijatrijskoj populaciji, razvijen je model za predikciju kompliciranih oblika apendicitisa. Zadržavajući sve ostale parametre identičnima kao u prethodno opisanom pristupu detekcije akutnog apendicitisa u odraslih, spojeni su negativni PHD nalazi i PHD nalazi nekompliciranog apendicitisa u jednu kategoriju, dok se druga sastojala od kompliciranih oblika bolesti.

Među napravljenim modelima najbolji se pokazao model slučajnih šuma s osjetljivošću od $0,997 \pm 0,001$ i specifičnošću $0,057 \pm 0,008$ (Tablica 13., Slika 25.).

Tablica 13. Performanse modela u predikciji kompliciranih oblika akutnog apendicitisa

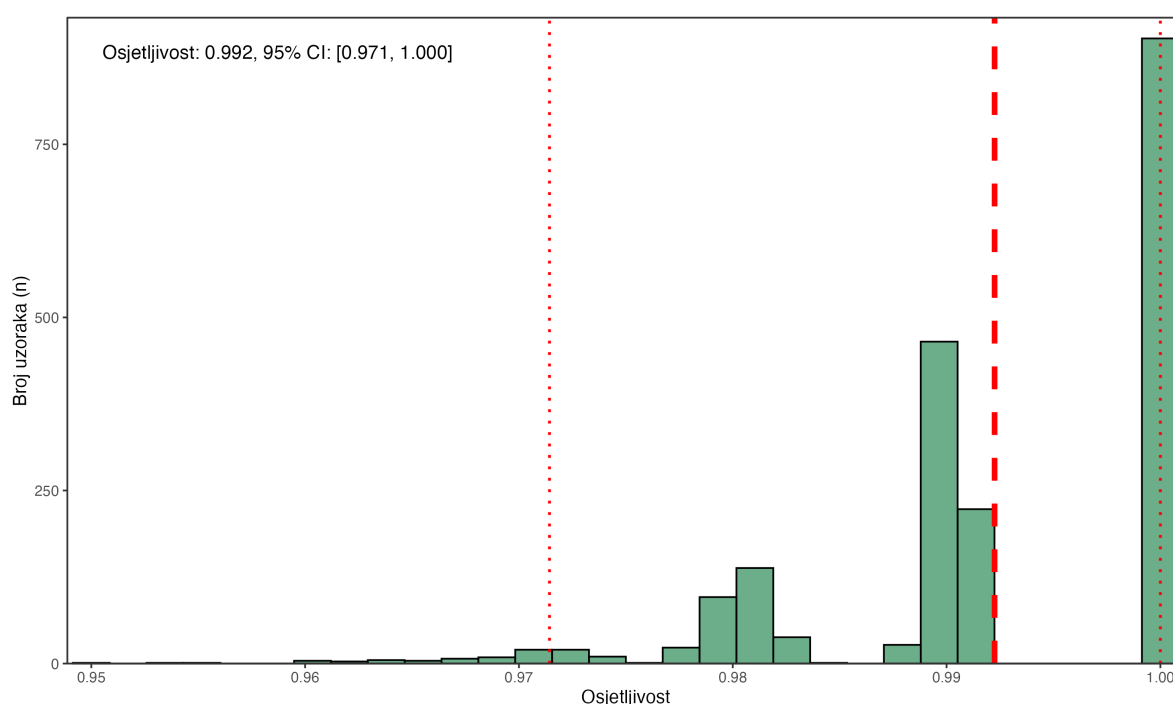
Model i prag odluke	AUC	Osjetljivost	Specifičnost
Slučajne šume			
PO1	0,785 ± 0,006	0,975 ± 0,003	0,151 ± 0,009
PO2		0,988 ± 0,002	0,091 ± 0,008
PO3		0,997 ± 0,001	0,057 ± 0,008
XGBoost			
PO1	0,781 ± 0,005	0,972 ± 0,003	0,205 ± 0,009
PO2		0,983 ± 0,002	0,141 ± 0,008
PO3		0,988 ± 0,004	0,108 ± 0,008
Logistička regresija			
PO1	0,776 ± 0,006	0,98 ± 0,003	0,108 ± 0,008
PO2		0,99 ± 0,002	0,074 ± 0,006
PO3		0,994 ± 0,002	0,053 ± 0,005

Podaci su prikazani kao prosjek ± standardna pogreška. Kratice: AUC: površina ispod ROC krivulje (engl. *Area under the curve*); PO1 – prag odluke pri kojem je osjetljivost minimalno 0,98, PO2 – prag odluke pri kojem je osjetljivost minimalno 0,99, PO3 – prag odluke pri kojem je osjetljivost >0,995

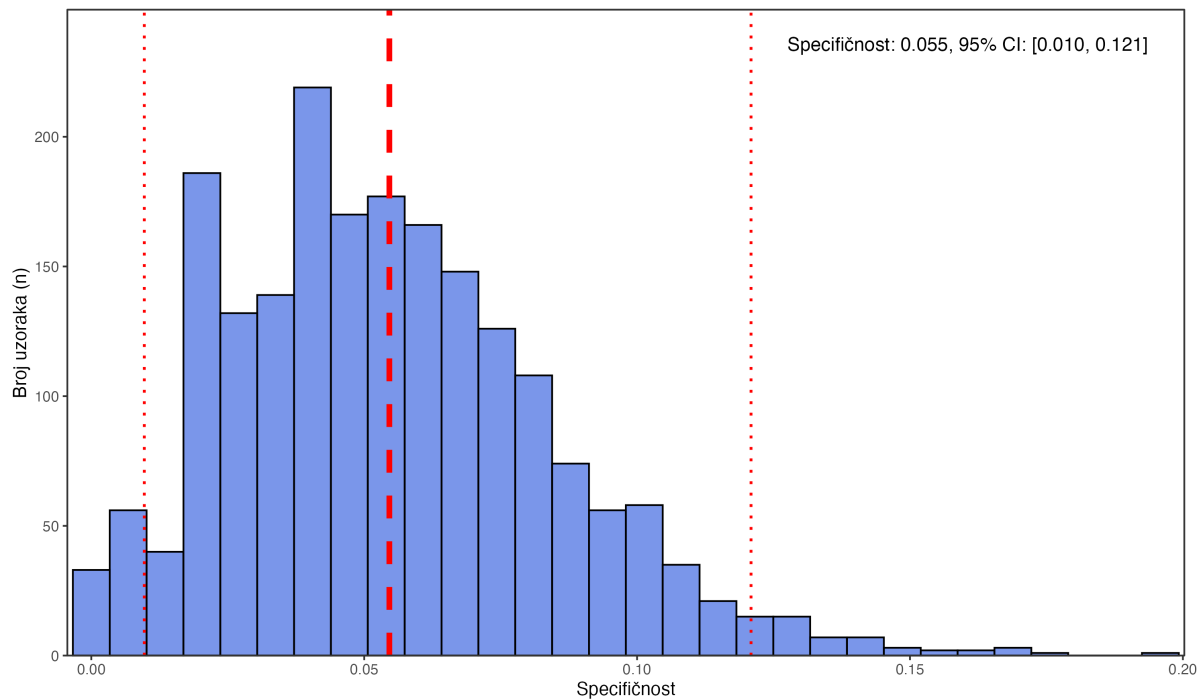


Slika 25. ROC krivulja za prosječnu vrijednost modela slučajnih šuma u predikciji kompliciranog oblika apendicitisa (plava krivulja) i ROC krivulje za svaki vanjski preklop (sive krivulje).

Ponovno su provedene dvije *bootstrap* analize odabranog modela – klasična i klaster *bootstrap*. U klasičnoj *bootstrap* analizi, za odabrani prag odluke PO3 slučajnih šuma, osjetljivost je bila 0,993 (95% CI 0,971 – 1), specifičnost 0,055 (95% CI 0,009 – 0,121), preciznost 0,531 (95% CI 0,471 – 0,596), PPV 0,52 (95% CI 0,462 – 0,582), NPV 0,888 (95% CI 0,603 – 1) i F1 ocjena 0,730 (95% CI 0,675 – 0,781) (**Slika 26.** i **27.**). U drugoj analizi, za odabrani prag odluke PO3, osjetljivost je bila 0,992 (95% CI 0,989 – 0,994), specifičnost 0,053 (95% CI 0,044 – 0,062), preciznost 0,521 (95% CI 0,518-0,523), PPV 0,521 (95% CI 0,518 – 0,523), NPV 0,869 (95% CI 0,822 – 0,912) i F1 ocjena 0,683 (95% CI 0,681 – 0,686).

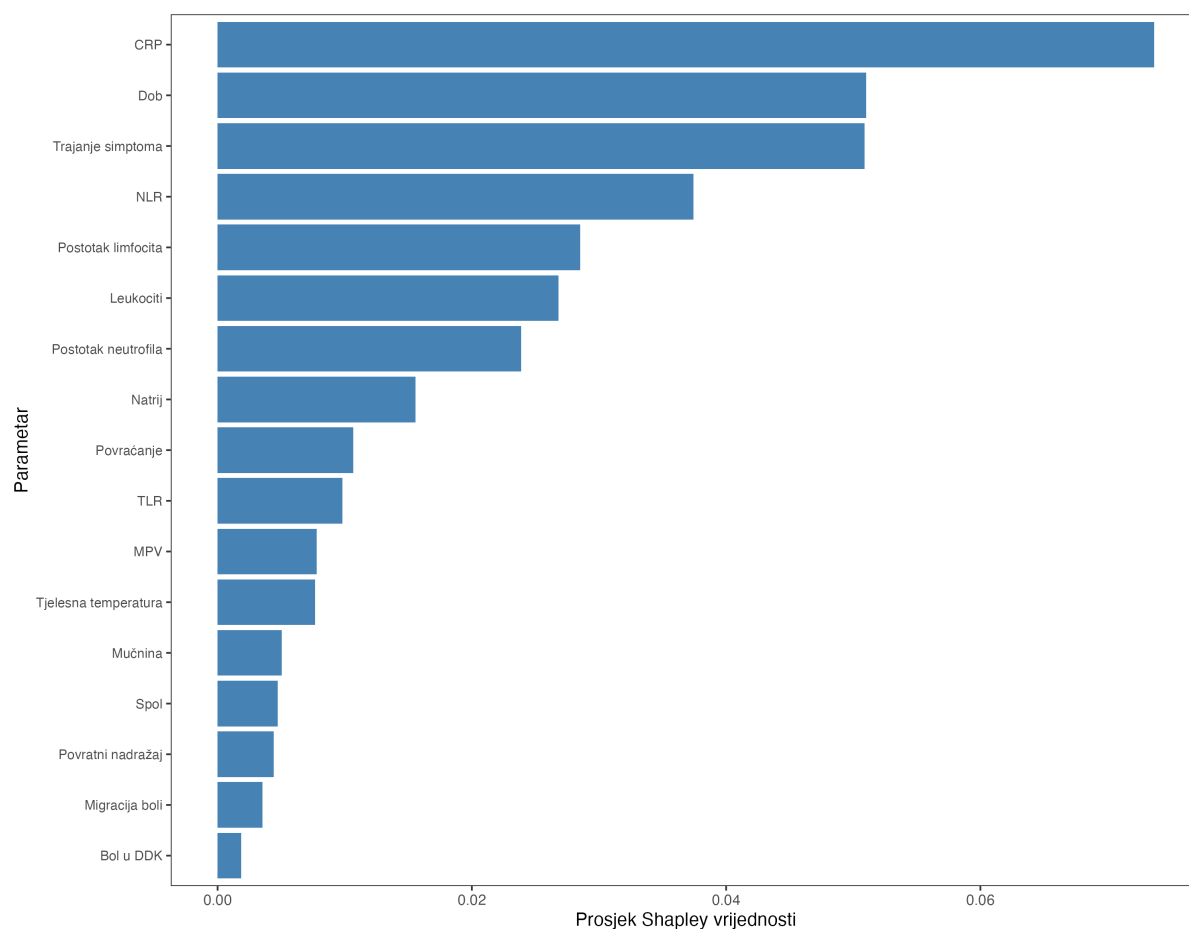


Slika 26. Prikaz prosječne osjetljivosti i 95% CI dobiven klasičnom *bootstrap* validacijom s 2000 iteracija za model slučajnih šuma na pragu odluke za koji je osjetljivost >0,995

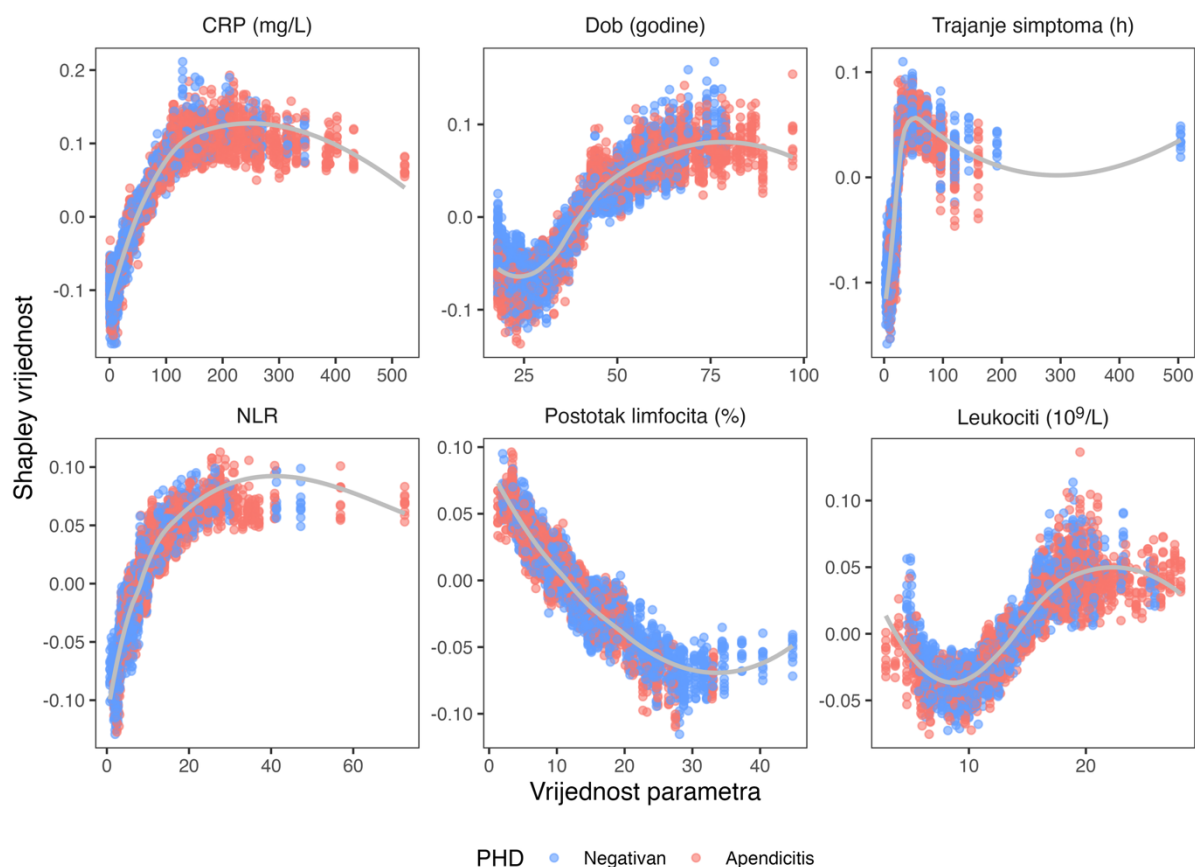


Slika 27. Prikaz prosječne specifičnosti i 95% CI dobiven klasičnom *bootstrap* validacijom s 2000 iteracija za model slučajnih šuma na pragu odluke za koji je osjetljivost $>0,995$

Ponovno, kako bismo bolje razumjeli kako je model obradio informacije, izračunate su Shapleyjeve vrijednosti prema odabranom pragu odluke PO3 za koji je osjetljivost $>0,995$. Provedena je SHAP analiza s 10000 permutacija izračunatih na *out-of-sample* uzorcima na odabranom pragu odluke (**Slika 28.**). Također, na **Slici 29.** prikazan je odnos najvažnijih numeričkih prediktora i SHAP vrijednosti.



Slika 28. Prikaz važnosti pojedinih značajni modela slučajnih šuma prema Shapleyjevim vrijednostima za odabrani prag odluke za koji je osjetljivost $>0,995$. Prikazano kao prosjek modula Shapley vrijednosti za svaki preklap. Kratice: DDK: donji desni abdominalni kvadrant; MPV: prosječni volumen eritrocita (engl. *mean platelet volume*); NRL: omjer neutrofila i limfocita (engl. *neutrophil to lymphocyte ratio*); TLR: omjer trombocita i limfocita (engl. *thrombocyte to lymphocyte ratio*)



Slika 29. Odnos najvažnijih prediktora i PHD nalaza modela slučajnih šuma prema Shapley vrijednostima za prag odluke za koji je osjetljivost $>0,995$. Kratice: NLR: omjer neutrofila i limfocita (engl. *neutrophil to lymphocyte ratio*); TLR: omjer trombocita i limfocita (engl. *thrombocyte to lymphocyte ratio*)

5. RASPRAVA

Akutni apendicitis predstavlja često hitno kirurško stanje, a u posljednje je vrijeme zabilježen porast istraživanja koja primjenjuju ML i AI radi unapređenja dijagnostike i liječenja. Suvremeni AI modeli – od klasičnih algoritama poput logističke regresije i neuronskih mreža do naprednog dubokog učenja – postigli su visoku razinu točnosti u prepoznavanju apendicitisa i razlikovanju stupnja bolesti (151).

Prvi dio istraživanja je usmjeren na pedijatrijsku populaciju, a drugi dio je proveden na podacima dobivenim iz odrasle populacije. U obje kohorte, podaci su prikupljeni iz populacije koja je podvrgnuta operacijskom liječenju zbog sumnje na akutni apendicitis. Takav pristup u startu označava populaciju visokorizičnu za postojanje apendicitisa. S obzirom na to da su svi pacijenti operirani, ne postoji uzorak stvarno negativnih ispitanika (neoperirani s negativnim PHD nalazom) i lažno negativni (neoperirani s pozitivnim PHD nalazom), već stvarno pozitivni (operirani s pozitivnim PHD nalazom), lažno pozitivni (operirani s negativnim PHD nalazom) što je bitno imati na umu kod interpretacije rezultata. Ovaj pristup se razlikuje od nekih provedenih studija koji uključuju neoperirane bolesnike kod kojih nedostaje PHD potvrda dijagnoze (141, 148, 152, 153). Usprkos tome, do sada su provedena i istraživanja koja su koristila pristup samo operiranih bolesnika (136, 139, 142). Pojedini autori pokušali su riješiti problem lažno negativne skupine kontaktiranjem bolesnika otpuštenih na kućnu njegu kako bi adresirali problem lažno negativne skupine, ali i dalje ostaje neadresiran problem stvarno negativnih slučajeva (143). Ovaj izbor ispitivane populacije temelji se na stavu da nije u potpunosti opravdano koristiti populaciju neoperiranih bolesnika za koje se pretpostavlja da boluju od akutnog apendicitisa u izradi ovakvih modela. Bez jasne patohistološke potvrde ne može se sa sigurnošću tvrditi da se radilo o akutnom apendicitisu bez obzira na provedenu obradu u smislu prikupljenih anamnestičkih podataka, kliničkog pregleda, učinjene laboratorijske te radiološke obrade što bi moglo dovesti do otežane interpretacije rezultata pogotovo uzimajući u obzir stopu negativnih apendektomija kao i varijabilne osjetljivosti i specifičnosti radioloških metoda (84, 154). Također, korištenje populacije općenitih bolova u trbuhu kao kontrolne skupine u smislu stvarno negativnih ispitanika nije u potpunosti opravdano. Kod takve populacije mogu nedostajati neke od najbitnijih značajki apendicitisa kao što su bolovi u donjem desnom kvadrantu, povišeni upalni parametri, pojava boli prije povraćanja, inapetencija i drugi, a temelju kojih bi uopće barem posumnjali na akutni apendicitis (155).

Rezultati u pedijatrijskoj populaciji su prikazali ML model razvijen primarno s ciljem smanjenja broja negativnih apendektomija u pedijatrijskoj populaciji. Stoga je bilo potrebno razviti model sa sposobnošću gotovo stopostotne detekcije stvarno pozitivnih slučajeva

apendicitisa kako bismo bili što sigurniji da, kada model daje negativnu dijagnozu ona nije lažno negativna već stvarno negativna. Takav pristup smo zadržali i za odraslu populaciju, uz prošireni prikaz rezultata prema različitim pragovima odluke. Od razvijenih modela u pedijatrijskoj populaciji, slučajne šume pokazale su najbolje performanse s osjetljivošću od $0,997 \pm 0,001$ i specifičnošću od $0,17 \pm 0,01$, dok je u odrasloj populaciji logistička regresija pokazala najbolju osjetljivost od $0,998 \pm 0,001$, uz lošiju specifičnost od $0,038 \pm 0,009$ na pragu odluke pri kojem je osjetljivost $>0,995$. Međutim, pri smanjenju praga odluke za koji je osjetljivost minimalno 0,98, XGBoost je pokazao osjetljivost od $0,972 \pm 0,004$ i specifičnost od 0,146. Modeli su trenirani na podacima bolesnika kod kojih je, temeljem kliničke procjene, već postojala visoka sumnja na apendicitis te su svi bili podvrgnuti kirurškom zahvatu. U skupu podataka nije bilo pacijenata koji nisu bili operirani, pa model nije mogao „naučiti“ prepoznati „prave negativne“ (bez apendicitisa, bez operacije) niti „lažno negativne“ (slučajevi apendicitisa koje je model propustio). Ovaj je model namijenjen pacijentima za koje bi kirurzi, zbog visoke sumnje na apendicitis, uobičajeno preporučili operaciju. U pedijatrijskoj populaciji model bi propustio apendicitis u samo 3 od 1000 slučajeva, a u odrasloj u 2 od 1000 slučajeva. Za usporedbu, kirurzi bi, oslanjajući se isključivo na standardne dijagnostičke modalitete, operirali svih 1000 pacijenata visokog rizika te bi ispravno identificirali sve slučajeve apendicitisa, ali bi u 8,5% slučajeva pogriješili u procjeni potrebe za operacijom u pedijatrijskoj populaciji i 10,3% u odrasloj populaciji (specifičnost 0%). Model slučajnih šuma u pedijatrijskoj populaciji postigao je specifičnost od 0,17, što znači da bi kod svakih 1000 bolesnika u ovako visokorizičnoj populaciji na kojoj je treniran, spriječio 14-15 nepotrebnih operacija djece koja zapravo nisu imala apendicitis, a pritom bi gotovo u potpunosti prepoznao prave slučajeve apendicitisa. U uzorku na kojem je stvoren model, 8 od 47 bolesnika s negativnim PHD nalazom bi izbjeglo operaciju. Model logističke regresije u odrasloj populaciji je postigao specifičnost od 0,038, što bi značilo da bi u ovako visokorizičnoj populaciji na kojoj je treniran spriječio 4 nepotrebne operacije na 1000 bolesnika koja zapravo nisu imala apendicitis. Konkretno, u uzorku na kojem je stvoren model, 2 od 56 bolesnika bi izbjegla operaciju. Međutim, XGB s pomakom praga na minimalnu osjetljivost od 0,98, spriječio bi 15 negativnih apendektomija na 1000 bolesnika, a na ovom uzorku 8 od 56. Takvi rezultati dolaze nauštrb osjetljivosti pa bi model propustio prepoznati 25 slučajeva apendicitisa na 1000 bolesnika. Ukratko, za bolesnike s već visokom sumnjom na apendicitis, naši bi model mogao pomoći u smanjenju broja nepotrebnih operacija u usporedbi s kliničkom procjenom, uz minimalan broj propuštenih slučajeva stvarnog apendicitisa.

Druga provedena istraživanja su također koristila ML za razvoj vlastitih modela u dijagnozi apendicitisa. Općenito, predstavljeni rezultati nadmašuju druga istraživanja u pogledu osjetljivosti, dok zaostaju u pogledu specifičnosti stoga bi trebalo usporediti naš rad s drugima u pogledu kompromisa između osjetljivosti i specifičnosti, pogotovo u kontekstu izbora populacije, broja ispitanika, izabranih parametara te same metodologije. Mijwil i Aggarwal su retrospektivno prikupili podatke za 625 ispitanika bez podjele na pedijatrijske i odrasle bolesnike, te navode osjetljivost 81,01% i specifičnost od 81,08% za model slučajnih šuma (133). Ovakav balansirani pristup dostiže dobru osjetljivost, ali bi propustio prepoznati 19% slučajeva apendicitisa. Međutim, njihov skup podataka se sastojao od operiranih i neoperiranih bolesnika, stoga i ova razlika osjetljivosti djelomično proizlazi iz sastava skupa podataka jer je njihov model treniran na općenitoj populaciji. Sličan pristup su primijenili Hsieh i sur., gdje su na populaciji od 180 ispitanika starijih od 16 godina testirali više modela, a kao najbolji su se pokazale slučajne šume s osjetljivosti od 0,94 i specifičnosti od 1. Iako su rezultati impresivni, njihov skup se sastojao od malenog broja bolesnika za koje je tražena kirurška konzultacija od kojih su neki otpušteni na kućnu njegu, a drugi operirani, korištena je metoda izdvajanja (engl. *holdout method*) se procjena vrši samo na jednom uzorku što je značajna razlika od ugniježdene unakrsne validacije. Za napomenuti je kako su se bavili problemom lažno negativne skupine tako što su telefonski kontaktirali otpuštene bolesnike kako bi provjerili ispravnost odluke kirurga (143). Akmese i sur. koristili različite algoritme na 595 pedijatrijskih bolesnika te postigli oko 93% osjetljivosti i 97 % specifičnosti primjenom stabala s podizanjem gradijenta (engl. *Gradient boosted trees*) koja je sadržavala i operirane i neoperirane bolesnike (141). Reismann i sur. predstavili su modele razvijene na 590 pedijatrijskih bolesnika s PHD potvrdom dijagnoze. Za razvoj modela u svrhu detekcije apendicitisa, koristili su se vrijednostima samo 5 biomarkera (leukociti, bazofili, monociti, neutrofil, CRP) i UZV promjerom apendicitisa čime su dobili osjetljivost od 0,93 i specifičnost od 0,67. Ponovno, korištena je metoda izdvajanja umjesto robusnijih metoda, skup podataka se sastojao od skoro 20% negativnih apendektomija što ga čini jako visokim, pogotovo u kontekstu što su ispitanici imali napravljen UZV abdomena preoperativno. Nadalje, vrijednost UZV promjera apendicitisa doprinosi 0,11 AUC vrijednosti u usporedbi s razvijenim modelom bez te vrijednosti (0,91 vs 0,8). Sve zajedno doprinosi manjoj nebalansiranosti skupa podataka kao i postojanju vrijednosti koja zasebno ima veliki značaj, a pogotovo uzimajući u obzir da je povećan promjer apendicitisa jedan od UZV parametara u dijagnosticiranju apendicitisa u konačnici i visokim vrijednostima osjetljivosti i specifičnosti (136). Marcinkevics i sur. napravili su sveobuhvatni pristup na pedijatrijskoj populaciji, koristeći javno dostupnu bazu

podataka hospitalizirane djece pod sumnjom na akutni apendicitis, od kojih je dio operiran te su imali PHD analizu, dok je dio liječen konzervativno, a dijagnoza se postavljala koristeći ultrazvučne kriterije skupa s kliničkim sustavima bodovanja (Alvarado i PAS). Svojim metodama dobili su osjetljivost 75-93% i specifičnost od 64-90% ovisno o modelu i kombinaciji korištenih parametara (148). Nadalje, Kendall i sur. koristili su kasnije istu bazu podataka s proširenim uzorkom od 781 pedijatrijskog bolesnika, te izvijestili o osjetljivosti 55-98%, specifičnosti 41-97% ovisno o modelu, korištenim parametrima i načinu validacije (156). Osim njih, Navaei i sur. isto su koristili navedenu bazu s podataka s uzorkom od 782 bolesnika, te izvijestili o točnosti do 99% te osjetljivosti do 100% i specifičnosti do 99% (153). Ova tri navedena istraživanja, iako su imala na raspolaganju jednake parametre koristila su različite modele i broj ispitanika. Također, Chadaga i sur. su koristili istu bazu podataka, ali su koristili ansambl modela i razvili svoj APPSTACK model te su koristili različite metode optimizacije (157). Međutim, Akbulut i sur. su u većoj seriji od 1797 isključivo operiranih bolesnika s PHD potvrdom, ne navodeći jasno je li se uzorak sastojao od pedijatrijskih bolesnika, i višom stopom negativnih apendektomija od 18,5%, koristeći CatBoost model izvijestili o osjetljivosti od 84% i specifičnosti od 93% (134).

Objektivna usporedba ovih istraživanja i njihovih modela s predstavljenim rezultatima je otežana jer su svi navedeni modeli koristili su različit broj kliničkih, laboratorijskih parametara, a neki su koristili i rezultate radioloških metoda, osim već navedenih istraživanja Marcinkevicsa i sur., Kendalla i sur., Navaeia i sur. te Chadaga i sur. koji su imala na raspolaganju jednake parametre uz različit broj ispitanika (148, 153, 156, 157). Važno je naglasiti da je niša našeg modela drukčija: on je namijenjen kohorti pacijenata s visokom sumnjom na apendicitis, kao završna „sigurnosna provjera“, a ne za probir svih pacijenata s abdominalnom boli. Najbolje bi se moglo usporediti s onim od Akbuluta i sur., ponajviše po odabiru ispitanika i parametrima, međutim imali su značajno višu stopu negativnih apendektomija te je korištena SMOTE (engl. *Synthetic minority oversampling technique*) metoda za rješavanje problema nebalansiranog skupa podataka uz nejasan konačan omjer apendicitisa i negativnih slučajeva što sve skupa dovodi do veće vrijednosti specifičnosti (134). Takvi modeli s umjetno balansiranim skupom podataka poboljšavaju ukupnu točnost i ostale performanse modela, ali inherentno dovode do povećane stope lažno pozitivnih nalaza i utječu na klasifikacijske performanse što je naš pristup nastojao svesti na minimum (158). Za razliku od njih, Erman i sur. su u razvoju svojih modela za predviđanje stupnja perforacije apendicitisa, bazirano na intraoperacijskom nalazu, jasno naveli stupanj povećanja uzorka za svaki model (159). Pri usporedbi studija valja imati na umu da nije dovoljno izolirano promatrati izvještene

mjere, nego i način na koji je model razvijen, podešen i validiran što je u ovom radu detaljno opisano.

Bitno je naglasiti da je u modelu slučajnih šuma u pedijatrijskoj populaciji veća vjerojatnost da su lažno negativne predikcije slučajevi nekomplikiranog oblika bolesti nego komplicirani oblici bolesti, što je očito iz izrazito visoke osjetljivosti za prepoznavanje gangrenoznog oblika od 0,9996 u ovom modelu. To je osobito relevantno jer neki autori predlažu da se samo komplicirani oblici apendicitisa trebaju odmah kirurški liječiti, dok se nekomplikirani slučajevi mogu inicijalno liječiti konzervativno i antibioticima. No, otprilike 20-30% tih slučajeva na kraju ipak bude podvrgnuto operacijskom liječenju (160, 161). Stoga, ako se u bolesnika visokog rizika s negativnim rezultatom ML modela primijeni antibiotska terapija s ekspektativnim pristupom, osjetljivost bi se efektivno mogla dodatno povećati smanjujući broj lažno negativnih.

Na temelju ovakvog pristupa testiran je dodatni klasifikacijski model u kojem su negativne apendektomije objedinjene s nekomplikiranim slučajevima te testirane u odnosu na komplicirane. Osim same dijagnoze, ključno je pitanje stratifikacije stupnja bolesti: predviđanje hoće li apendicitis biti nekomplikiran ili je već napredovao u komplicirani stupanj. To može imati terapijske implikacije u ustanovama koje provode i konzervativno liječenje. Zadržavajući sve ostale parametre jednakima kao u gore opisanom pristupu, modelom slučajnih šuma postigli smo usporedivu osjetljivost 0,994 i nešto nižu specifičnost 0,129 u pedijatrijskoj populaciji, dok je u odrasloj populaciji osjetljivost također bila tek neznatno niža 0,997, a specifičnost je porasla 0,057. Nadalje, na PO3, XGB je pokazao nižu osjetljivost od 0,972 i višu specifičnost od 0,205. Teoretski, takav bi se model mogao primijeniti pri odabiru alternativnog pristupa liječenju apendicitisa, u kojem bi samo pacijenti s kompliciranim apendicitisom bili odmah kirurški liječeni, dok bi se nekomplikirani u početku pratili uz primjenu antibiotske terapije. Ovaj bi pristup povećao broj lažno negativnih slučajeva pošteđenih od operacije, po cijenu dodjeljivanja negativnih predikcija nekomplikiranim apendicitisima. Međutim, budući da su krivo klasificirani nekomplikirani apendicitisi imali najniže vjerojatnosti za apendicitis prema modelu, udio pacijenata koje je model označio kao negativne, a koji bi nakon konzervativnog liječenja ipak morali na operaciju, vjerojatno bi bio manji od pretpostavljenih 20-30%. U tom je smislu vrijedno naglasiti da sve veći broj dokaza sugerira kako svi pacijenti s apendicitisom ne napreduju prema perforaciji te da se perforacija pretežito događa prije dolaska pacijenata u bolnicu (62, 162). Nadalje, čini se da je stopa smrtnosti nakon perforacije niža, a nakon negativne apendektomije viša nego što se ranije smatralo, što naglašava važnost ispravne dijagnoze u odnosu na brzu (163).

Phan-mai i sur. su proveli retrospektivno istraživanje na 1950 ispitanika s PHD potvrdom bez podjele na pedijatrijske i odrasle s ciljem detekcije kompliciranih oblika apendicitisa. Koristili su više klasifikacijskih metoda, uključujući i ANN s više korištenih parametara po modelu, a osim toga izvijestili su svoje rezultate i bez balansiranja baze podataka kao i rezultate dobivene koristeći SMOTE. Dobili su vrijednosti točnosti 0,703 do 0,81 na originalnom nebalansiranom setu, i točnost od 0,655 do 0,82 koristeći SMOTE (142). Kang i sur. su razvili model na 136 odraslih ispitanika s PHD potvrdom koristeći standardne biokemijske markere, kliničke podatke te imunohistokemijske markere te izvijestili o visokim vrijednostima osjetljivosti od 0,75 do 0,93 i specifičnosti od 0,645 do 1 ovisno o uspoređivanim skupinama. No, ukupan broj ispitanika od 127 dodatno podijeljen na nebalansirane skupine od 23 i 104, značajno limitira njihove rezultate (147). Marcinkevics je, kao što je učinjeno u prezentiranom radu, osim predikcije dijagnoze napravio i analizu predikcije stupnja apendicitisa na istoj populaciji. Izvijestili su o osjetljivosti od 0,93 do 0,98 i specifičnosti od 0,37 do 0,52. U svojoj usporedbi jednu skupinu je sačinjavala populacija s kompliciranim apendicitisima, a drugu nekomplicirani ili oni bez upale. Međutim, njihov model ponovno nije imao definitivnu PHD dijagnozu za sve bolesnike. Samo oni s kompliciranim apendicitisima su imali jer su bili podvrgnuti kirurškom liječenju, dok nekomplicirani i oni bez upale nisu uvijek imali PHD potvrdu jer nisu bili operirani (148). Wei i sur. su na populaciji starijih od 65 napravili više ML modela za razlikovanje nekompliciranog od kompliciranog apendicitisa, od kojih se GBM pokazao kao najuspješniji s točnosti od 0,956, osjetljivosti od 0,917 i specifičnosti od 0,973 (164).

Isto kao i u slučaju modela za predikciju apendicitisa, gore navedeni modeli razvijeni za predikciju nekompliciranih i kompliciranih oblika bolesti su razvijeni koristeći različit broj kliničkih, laboratorijskih parametara, dok su neki također koristili i rezultate radioloških metoda. Također, modeli nisu razvijeni na metodološki jednak način što dodatno otežava usporedbu rezultata.

Provedena su dva pristupa bootstrap validacije za procjenu robusnosti modela. Klasična bootstrap analiza procjenjuje varijabilnost performansi pri ponovnom učenju na različitim podskupovima iste populacije, što procjenjuje stabilnost modela u odnosu na fluktuacije uzorkovanja kojom je u odrasloj populaciji osjetljivost 0,991 i specifičnost 0,049 za predikciju apendicitisa, dok je model za razlikovanje nekompliciranog od kompliciranog oblika bolesti pokazao osjetljivost od 0,991 i specifičnost od 0,049. Klaster *bootstrap* analiza vanjskih preklopa iz ugniježdene unakrsne validacije reflektira varijabilnost koja proizlazi iz različitih mogućih podjela podataka na trening i test setova kojom se dobila osjetljivost od 0,992 i

specifičnost od 0,043 za model za predikciju apendicitisa dok je model za razlikovanje nekomplikiranog od kompliciranog oblika bolesti imao osjetljivost od 0,992 i 0,053. Takvi rezultati pokazuju da su modeli minimalno osjetljiv na varijabilnost pri ponovnom uzorkovanju. Konzistentnost rezultata između oba *bootstrap* pristupa potvrđuje robusnost modela i pouzdanost procjene performansi, što je osobito važno za kliničku implementaciju gdje je stabilnost dijagnostičkih algoritma ključna za sigurnost pacijenata.

Analiza Shapleyjevih vrijednosti u našem istraživanju identificirala je CRP, leukocite, duljinu trajanja simptoma, postotak neutrofila i koncentraciju natrija kao najznačajnije prediktore u pedijatrijskoj populaciji za detekciju apendicitisa, dok su za razlikovanje nekomplikiranog od kompliciranog oblika bolesti kao najznačajniji identificirani CRP, koncentracija natrija, NLR, postotak limfocita i trajanje simptoma. U odrasloj populaciji za detekciju apendicitisa kao najznačajniji prediktori identificirani su postotak neutrofila, NLR, TLR, spol i postotak limfocita, dok su za razlikovanje nekomplikiranog od kompliciranog oblika bolesti identificirani CRP, dob, duljina trajanja simptoma, NLR i postotak limfocita. Ovi rezultati se djelomično podudaraju s rezultatima drugih istraživanja koja su koristila Shapley vrijednosti za objašnjivost ML modela u detekciji apendicitisa. Isto tako, osim općenite SHAP vrijednosti po parametrima, pojedina istraživanja su koristila dodatne prikaze SHAP rezultata kao što je npr. *beeswarm plot*, čime se dodatno postiže objašnjivost rezultata u kontekstu doprinosna vrijednosti parametara pozitivnim ili negativnim predikcijama (134, 152, 164-166). Osim Shapleyjevih vrijednosti, korištene su i druge metode kao što su *mean decrease in accuracy*, *recursive feature elimination*, kombinaciju više metoda ili pak ne navodeći jasno metodu osim da je napravljena analiza važnosti parametara (148, 153, 167-169).

Rezultati modela za predikciju su uspoređeni s AIR sustavom bodovanja u pedijatrijskoj populaciji. AIR na bodovnom pragu od 0 na ovoj populaciji pokazuje nešto višu osjetljivost od 0,998 naspram 0,997, međutim specifičnost mu je 0. Već na bodovnom pragu od 1 pokazuje nižu osjetljivost od 0,979 dok mu je specifičnost od 0,17 komparabilna našoj. Za sve ostale pragove ima nižu osjetljivost te višu specifičnost. U srednjem rasponu (5–8 bodova) kompromis postaje još nepovoljniji za AIR. Na bodovnom pragu 5 osjetljivost pada na 0,31, dok je kod našeg modela i dalje približno 1, uz specifičnost od 0,92 za AIR skor i 0,17 za naš model. Drugim riječima, iako AIR skor postiže veću specifičnost, to je postignuto na račun gubitka sposobnosti prepoznavanja većine stvarnih slučajeva apendicitisa. Na bodovnom pragu 8, već i prije praga 9 za kategoriziranje za visoke vjerojatnosti za apendicitis, AIR ima osjetljivost 0 dok mu je specifičnost 1 što se zadržava za sve iduće bodovne pragove. Ovakvi rezultati proizlaze iz same konstrukcijske logike AIR sustava bodovanja, u kojem viši zbroj

bodova odražava veću vjerojatnost apendicitisa. Prema meta analizi Anderssona i Starka iz 2025. godine, na nižim pragovima (≤ 3 ili ≤ 4) postiže se vrlo visoka osjetljivost (0,95 i 0,99), budući da gotovo svi bolesnici s apendicitisom prelaze taj prag; međutim, specifičnost se time svodi na zanemarivu razinu, jer se u pozitivnu kategoriju svrstavaju i mnogi bez bolesti. Klinički gledano, to definira skupinu bolesnika niskog rizika s izrazito niskom prevalencijom kompliciranog apendicitisa (0,3–1%), kod kojih je opravdano odgoditi invazivnu dijagnostiku i provesti promatranje ili revalvaciju. Suprotno tome, visoki pragovi (≥ 9) postižu visoku specifičnost od 0,98, jer takve kliničko-laboratorijske vrijednosti obilježavaju bolesnike s apendicitisom. Ipak, osjetljivost u tom slučaju dramatično pada, budući da se samo oko četvrtine svih bolesnika s apendicitisom klasificira u visokorizičnu skupinu (170). Drugim riječima, AIR sustav bodovanja funkcionira primarno kao instrument za stratifikaciju rizika u tri kategorije, te se ne može koristiti kao dijagnostički model sposoban istodobno maksimizirati i osjetljivost i specifičnost. Što se tiče diskriminatorskih vrijednosti značajki korištenim u bodovanju AIR sustava bodovanja, zabilježene su sličnosti s razvijenim modelima pa se tako primjećuju vrijednosti CRP-a oko 10mg/dL, $10 \times 10^9/L$ za leukocite i 80% za postotak neutrofila (89). To dodatno pokazuje potencijalni značaj modela za fino ugađanje postojećih kliničkih sustava bodovanja za dijagnozu apendicitisa.

Recentno, Kucukuakcali i Cicek su napravili retrospektivnu evaluaciju RIPASA, ALVARADO i AIR sustava bodovanja koristeći XGBoost model navodeći osjetljivost od 0,978 i specifičnost od 0,294 za AIR prema konačnom PHD nalazu kada se gledalo je li bodovni sustav prepoznao apendicitis. Dodatno je korišten XGBoost kako bi se procijenila diskriminativna snaga ovih sustava bodovanja. Kao jedini parametar su koristili vrijednosti ovih bodovnih sustava, te je za AIR postignuta osjetljivost od 0,909 i specifičnost od 1 što bi značilo da je model na njihovoj populaciji našao prag na kojem bi mogao odlučiti radi li se o apendicitisu ili ne. Međutim, ova studija je provedena na uzorku od 107 bolesnika, te iako je provedena unakrsna validacija uzorak je nedostatan da bi rezultati bili od značaja kao i što nedostaje deskriptivne statistike u vidu karakteristika bolesnika koja bi pomogla u interpretaciji rezultata (171). Usprkos tome, ova studija pokazuje potencijalni značaj ML modela u razvoju alata za pomoć u odlučivanju koji se koriste i klasičnim sustavima bodovanja i ML modelima za daljnje fino ugađanje.

Uzimajući u obzir postojanje AI modela za korištenje na trijaži, potom u predikciji apendicitisa nakon kliničkog pregleda i laboratorijskih nalaza, radiomiku, intraoperacijsku potporu, predikciju stupnja bolesti prema PHD nalazu i predikciju postoperativnih komplikacija AI i ML modeli bi se mogli sveobuhvatno koristiti u dijagnostici apendicitisa

automatizirajući cijeli proces. Uz sve to potrebno je i razmatrati etičke i praktičke implikacije, prvenstveno kako bi se izgradilo povjerenje u ove sustave i minimizirao rizik pristranosti i pogrešaka kada se koriste u stvarnim kliničkim uvjetima (151). Jedan od istaknutijih primjera je onaj od Marcinkevicsa i sur. gdje su napravili mrežni Alat za predikciju pedijatrijskog apendicitisa kao prototip na temelju svoje studije (172). Implementacija ovih modela zahtijeva i odabir prikladnog praga odluke koji balansira osjetljivost i specifičnost i optimizira PPV i NPR ovisno o lokalnim zahtjevima i preferencijama, opterećenosti hitne službe, prevalenciji apendicitisa i organizaciji upućivanja na hitne prijeme iz primarne skrbi (152).

Izložena provedena istraživanja imaju određena ograničenja. Kao prvo, podaci korišteni za izradu modela su prikupljeni retrospektivno, stoga bi se trebala provesti i interna validacija na većem skupu podataka kako bi se bolje procijenili rezultati specifičnosti modela zbog malog broja negativnih apendektomija. Nadalje, prikupljeni su podaci iz samo jednog centra stoga bi se na tragu već provedenih jednocentričnih i multicentričnih istraživanja provesti i dodatna eksterna validacija modela u svrhu bolje generalizacije podataka i isključenja mogućnosti prenaučivosti modela kao razloga ovih rezultata (168, 173). Dio podataka korišten za izradu modela u pedijatrijskoj populaciji dobiven je imputacijom, ali je imputacija svedena na minimum u drugom dijelu istraživanja. Nadalje, ciljna varijabla u skupu podataka (PHD) nema normalnu distribuciju, nego je većina ispitivane populacije s patohistološkom dijagnozom apendicitisa, a tek manji dio s inoentnim crvuljkom na PHD-u što je očekivano uzimajući u obzir izbor populacije. Ovaj problem neuravnoteženosti ciljne varijable općenito je očekivan u istraživanjima o apendicitisu, gdje se stope negativnih apendektomija kreću između 1 i 40% ovisno o populaciji (174).

U perspektivi, problem nedostatka lažno negativnih i stvarno negativnih uzoraka u skupu podataka, mogao bi se djelomično riješiti daljnjim prikupljanjem podataka, barem na populaciji bolesnika koja je bila ili će biti pregledana u KBC Split. Naime, praksa je u ovoj ustanovi dio bolesnika s bolovima u trbuhu naručiti na kontrolni pregled idući dan ili u slučaju pogoršanja stanja i ranije. Tako da je zasigurno dio od ispitivane populacije, dan ranije bio na prvom pregledu kada nije postavljena sumnja na operacijsko liječenje ili hospitalizaciju, već na daljnje praćenje. Dostupni podaci iz tog prvog pregleda bi se potencijalno mogli iskoristiti kao uzorak za lažno negativnu skupinu, ako je nakon provedenog operacijskog liječenja od idućeg dana PHD nalaz bio pozitivan. Nasuprot tome, problem skupine stvarno negativnih teže je rješiv. Ako bismo uzeli bolesnike s bolovima u donjem desnom kvadrantu trbuha i imali sumnju na mogući razvoj akutnog apendicitisa te ih naručili na kontrolu idući dan, ostalo bi pitanje spontanog razrješenja inicijalnog upalnog zbivanja ili je pak uzrok boli neki drugi uzrok

u slučaju urednih laboratorijskih nalaza. Također, bolesnicima s porastom upalnih parametara na kontrolnom pregledu često se napravi dodatna obrada, u smislu dodatnih radioloških pretraga ili konzilijarnih pregleda pa se u kontrolnom intervalu uspije postaviti konačna dijagnoza ili se isključi akutni apendicitis. U konačnici, problem stvarno negativne skupine leži u tome koju bismo točno kombinaciju anamnestičkih i kliničkih podataka trebali koristiti u postavljanju sumnje na akutni apendicitis, a ne oslanjajući se na iskustveno znanje.

Nadalje, problem neuravnoteženosti ciljne varijable bi se mogao riješiti provedbom multicentričnih istraživanja kako bi se dobio podjednak broj ispitanika u svakoj skupini. Odnosno, prikupljenom broju inoentnih crvuljaka trebalo bi pridružiti podjednak broj apendicitisa, ili nekomplikiranih te komplikiranih apendicitisa, koji bi bili usklađeni po ulaznim varijablama. Metodama poput SMOTE-a umjetno se povećava broj ispitanika u manjinskoj skupini što može poboljšati generalizacijsku sposobnost modela, ali i potencijalno dovesti do prenaučenosti modela. Osim toga, zbog linearne interpolacije uzorci mogu odstupati od temeljne distribucije podataka (174). Također, neke tehnike ponovnog uzorkovanja poput poduzorkovanja (engl. *undersampling*), klasterkih tehnika kao i metoda augmentacije podataka, metoda ansambla, troškovno osjetljivog (engl. *cost sensitive*) učenja te strategija inženjeringa (engl. *feature engineering*) i selekcije značajki bi mogle poboljšati ciljne metrike kod nebalansiranih skupova podataka uz pažljivu evaluaciju (175).

Zaključno, visoka osjetljivost ovih modela čini ih posebno vrijednima u kontekstu brze procjene rizika u visokorizičnoj populaciji za postojanje akutnog apendicitisa nakon napravljene minimalne obrade – anamneza, klinički pregled i osnovni laboratorijski nalazi. Doprinos je u tome što se pokazuje kako se usmjeravanje na cilj bez propuštanja stvarno pozitivnih slučajeva, i korištenjem odgovarajućeg skupa podataka bolesnika visokog rizika, može izgraditi model kojim bi se za dio bolesnika moglo izbjeći nepotrebno operacijsko liječenje. Ovakvi rezultati naglašavaju da ne postoji univerzalna metrika primjenjiva na sve ML modele za apendicitis, već mora biti usklađena s predviđenom svrhom upotrebe. Za probirni alat u hitnim službama mogla bi se prihvatiti niža osjetljivost ako se time znatno poboljšava specifičnost. No za alat koji podržava donošenje odluka u trenutku kirurške intervencije, što je naš slučaj uporabe, maksimalna osjetljivost je od presudne važnosti. Ovim se postiže razina sigurnosti, a istodobno pruža i klinička korist u vidu smanjenja broja nepotrebnih operacija. Buduća istraživanja bi se trebala fokusirati na multicentrični ustroj uz multicentrične eksterne validacije uz jasno specificiranje kojoj je kliničkoj ulozi ML model namijenjen – potvrđivanju ili isključivanju dijagnoze, jer upravo to temeljno određuje kako bi model trebao biti evaluiran. Ovaj rad pruža referentnu točku za pristup usmjeren na

isključivanje i pokazuje da je moguće postići gotovo savršenu osjetljivost uz nenultu i zadovoljavajuću specifičnost. Daljnjim razvojem, ovi modeli bi se mogli koristiti kao korisne potpore u svakodnevnoj kliničkoj praksi putem integriranja u bolničke informatičke sustave ili razvojem kao samostalni digitalni alati za procjenu rizika.

6. ZAKLJUČCI

1. U pedijatrijskoj populaciji model slučajnih šuma postigao je osjetljivost od 0,997 i specifičnost 0,17 u predikciji akutnog apendicitisa
2. U pedijatrijskoj populaciji model slučajnih šuma postigao je osjetljivost od 0,994 i specifičnost od 0,129 u predikciji kompliciranog oblika bolesti
3. U odrasloj populaciji logistička regresija postigla je osjetljivost od 0,998 i specifičnost od 0,038 u predikciji akutnog apendicitisa
4. U odrasloj populaciji model slučajnih šuma postigao je osjetljivost od 0,997 i specifičnost od 0,057 u predikciji kompliciranog oblika bolesti
5. Razvijeni model slučajnih šuma u pedijatrijskoj populaciji nadmašuje AIR sustav bodovanja u većini točaka ROC krivulje
6. SHAP analiza pokazala je CRP, leukocite, duljinu trajanja simptoma, postotak neutrofila i koncentraciju natrija kao parametre od najvećeg značaja za slučajne šume u detekciji akutnog apendicitisa u pedijatrijskoj populaciji
7. SHAP analiza pokazala je CRP, koncentraciju natrija, omjer neutrofila i limfocita, postotak limfocita i duljinu trajanja simptoma kao parametre od najvećeg značaja za slučajne šume u detekciji kompliciranog oblika apendicitisa u pedijatrijskoj populaciji
8. SHAP analiza pokazala je postotak neutrofila, omjer neutrofila i limfocita, omjer trombocita i limfocita, spol i postotak limfocita kao parametre od najvećeg značaja za logističku regresiju u detekciji akutnog apendicitisa u odrasloj populaciji
9. SHAP analiza pokazala je kao CRP, dob, duljinu trajanja simptoma, omjer neutrofila i limfocita i postotak limfocita parametre od najvećeg značaja za slučajne šume u detekciji kompliciranog oblika apendicitisa u odrasloj populaciji
10. *Bootstrap* validacija potvrdila je performanse modela u odrasloj populaciji s uskim intervalima pouzdanosti kroz različite uvjete uzorkovanja

Istraživanje pokazuje da je, na kohorti bolesnika s već visokom kliničkom sumnjom, moguće izgraditi modele strojnog učenja koji uz gotovo savršenu osjetljivost ostvaruju nenultu i zadovoljavajuću specifičnost, čime se smanjuje broj negativnih apendektomija uz zanemariv rizik od propuštanja stvarno pozitivnih slučajeva. Postignute visoke osjetljivosti omogućavaju smanjenje rizika od propuštenih dijagnoza, dok objašnjivost modela putem SHAP analize osigurava transparentnost i jača kliničko povjerenje, a bootstrap validacije potvrđuju stabilnost rezultata. Ujedno, modeli za predikciju kompliciranih oblika apendicitisa također su pokazali vrlo visoku osjetljivost uz nenultu i zadovoljavajuću specifičnost, čime se otvara mogućnost njihove primjene u terapijskoj stratifikaciji bolesnika i ranijem prepoznavanju onih s težim kliničkim tijekom.

Unatoč ograničenjima, poput relativno niske specifičnosti i retrospektivnog dizajna, razvijeni modeli predstavljaju značajan korak naprijed prema personaliziranoj dijagnostici apendicitisa, osobito u populacijama gdje je klinička procjena otežana. Buduća istraživanja trebaju biti usmjerena na multicentričnu prospektivnu validaciju, fino ugađanje pragova odluke u skladu s lokalnom prevalencijom i kapacitetima sustava te integraciju dodatnih prediktora, uključujući radiološke parametre, uz konačnu implementaciju modela u bolničke informacijske sustave kao potporu donošenju odluka.

7. SAŽETAK

Uvod: Akutni apendicitis predstavlja najčešće hitno kirurško stanje u abdominalnoj kirurgiji, s globalnom incidencijom od približno 214 slučajeva na 100.000 stanovnika. Unatoč stalnim poboljšanjima u dijagnostici, stopa negativnih apendektomija još uvijek varira između 6,6% i 31,9% ovisno o populaciji. Modeli strojnog učenja (ML) pokazali su obećavajuće rezultate u medicinskim dijagnostičkim sustavima, uključujući potencijal za poboljšanje točnosti dijagnoze apendicitisa i smanjenje nepotrebnih kirurških zahvata.

Ciljevi: Glavni cilj istraživanja bio je razvoj objašnjivih modela strojnog učenja za predviđanje akutnog apendicitisa u pedijatrijskoj i odrasloj populaciji sa sumnjom na akutni apendicitis. Hipoteza je bila da će modeli strojnog učenja (slučajne šume i XGBoost) pokazati bolje dijagnostičke performanse od logističke regresije kao i do sada razvijenih kliničkih alata.

Ispitanici i postupci: Za pedijatrijsku populaciju retrospektivno su prikupljeni podaci za 551 dijete operirano u KBC Split (01/2019-06/2023), dok je za odraslu populaciju uključeno 542 bolesnika (01/2020-06/2024). Svi ispitanici imali su patohistološku potvrdu dijagnoze. Izrađena su tri ML modela: slučajne šume, XGBoost i logistička regresija. Korišten je pristup ugniježdene unakrsne validacije s petostrukom unutarnjom i vanjskom validacijom, ponovljenom 10 puta. Poseban naglasak stavljen je na postizanje maksimalne osjetljivosti kako bi se minimizirao broj propuštenih slučajeva apendicitisa. Objašnjivost modela postignuta je SHAP analizom.

Rezultati: U pedijatrijskoj populaciji od ukupno 551 ispitanika, 47 je imalo negativan PHD nalaz, 252 nekomplikirani i 252 komplicirani apendicitis. Model slučajnih šuma pokazao se najboljim za predikciju apendicitisa s osjetljivošću $0,997 \pm 0,001$ i specifičnošću $0,17 \pm 0,01$. Za razlikovanje kompliciranog od nekomplikiranog apendicitisa, isti model postigao je osjetljivost $0,994 \pm 0,002$ uz specifičnost $0,129 \pm 0,009$. U odrasloj populaciji od 542 bolesnika (56 negativnih, 210 nekomplikiranih, 276 kompliciranih), model logističke regresije postigao je najbolji AUC od $0,765 \pm 0,01$. Na pragu maksimalne osjetljivosti ($0,998 \pm 0,001$), specifičnost je iznosila $0,038 \pm 0,009$. Za razlikovanje stupnja bolesti, slučajne šume postigle su osjetljivost $0,997 \pm 0,001$ i specifičnost $0,057 \pm 0,008$. SHAP analiza identificirala je CRP, leukocyte, trajanje simptoma, postotak neutrofila i natrij kao najznačajnije prediktore u pedijatrijskoj populaciji, dok su u odrasloj populaciji to bili postotak neutrofila, NLR, TLR i spol.

Zaključak: Istraživanje je uspješno demonstriralo mogućnost razvoja ML modela s gotovo savršenom osjetljivošću za detekciju akutnog apendicitisa u visokorizičnoj populaciji. Modeli bi mogli smanjiti broj negativnih apendektomija za 14-15 slučajeva na 1000 operiranih u pedijatrijskoj populaciji, odnosno 4 slučaja na 1000 u odrasloj populaciji, uz minimalan rizik

propuštanja stvarnih slučajeva apendicitisa. Nadalje, ovi modeli bi se mogli koristiti i za detekciju kompliciranih oblika apendicitisa. Rezultati pružaju temelj za buduću implementaciju ML alata u kliničku praksu kao potporu odlučivanju u trenutku kirurške intervencije.

8. SUMMARY

TITLE:**Machine learning models in the diagnostics of acute appendicitis in children and adults**

Introduction: Acute appendicitis represents the most common emergency surgical condition in abdominal surgery, with a global incidence of approximately 214 cases per 100,000 population. Despite continuous improvements in diagnostics, negative appendectomy rates still vary between 6.6% and 31.9% depending on the population. Machine learning (ML) models have shown promising results in medical diagnostic systems, including potential for improving appendicitis diagnostic accuracy and reducing unnecessary surgical procedures.

Objectives: The main objective was to develop explainable machine learning models for predicting acute appendicitis in pediatric and adult populations with suspected acute appendicitis. The hypothesis was that ML models (random forests and XGBoost) would demonstrate superior diagnostic performance compared to logistic regression and previously developed clinical tools.

Patients and Methods: For the pediatric population, data were retrospectively collected for 551 children operated at University Hospital of Split (01/2019-06/2023), while for the adult population, 542 patients were included (01/2020-06/2024). All participants had histopathological confirmation of diagnosis. Three ML models were developed: random forests, XGBoost, and logistic regression. Nested cross-validation approach with five-fold inner and outer validation, repeated 10 times, was employed. Special emphasis was placed on achieving maximum sensitivity to minimize missed appendicitis cases. Model interpretability was achieved through SHAP analysis.

Results: In the pediatric population of 551 patients, 47 had negative histology, 252 uncomplicated, and 252 complicated appendicitis. Random forests model performed best for appendicitis prediction with sensitivity 0.997 ± 0.001 and specificity 0.17 ± 0.01 . For distinguishing complicated from uncomplicated appendicitis, the same model achieved sensitivity 0.994 ± 0.002 with specificity 0.129 ± 0.009 . In the adult population of 542 patients (56 negative, 210 uncomplicated, 276 complicated), logistic regression achieved the best AUC of 0.765 ± 0.01 . At maximum sensitivity threshold (0.998 ± 0.001), specificity was 0.038 ± 0.009 . For disease severity differentiation, random forests achieved sensitivity 0.997 ± 0.001 and specificity 0.057 ± 0.008 . SHAP analysis identified CRP, leukocytes, symptom duration, neutrophil percentage, and sodium as the most significant predictors in the pediatric population, while in adults these were neutrophil percentage, NLR, TLR, and sex.

Conclusion: The study successfully demonstrated the possibility of developing ML models with near-perfect sensitivity for detecting acute appendicitis in high-risk populations. The models could reduce negative appendectomies by 14-15 cases per 1000 operations in the pediatric population and 4 cases per 1000 in the adult population, with minimal risk of missing true appendicitis cases. Furthermore, these models could be used for detection of complicated appendicitis cases. The results provide a foundation for future implementation of ML tools in clinical practice as decision support at the point of surgical intervention.

9. LITERATURA

1. FIPAT. Terminologia Anatomica. International anatomical terminology. 2. izdanje. Georg Thieme Verlag. 2019.
2. Patra A, Kaur H, Chhabra U, Sghar A, Balawender K, Pasternak A i sur. Reappraisal of the variational anatomy of the vermiform appendix and their possible clinical applicability: a cadaveric analysis. *Folia Morphol (Warsz)*. 2023;83:716-726.
3. Glover J. The human vermiform appendix—a general surgeon's reflections, *CEN Tech. J.* 1988;3:31-8.
4. Lambert SW. Description of the vermiform appendix from the "de fabrica" of Vesalius. *Ann Med Hist.* 1937;9:422-7.
5. Collins DC. Historic phases of appendicitis. *Ann Surg.* 1931;94:179-96.
6. Herrod PJJ, Kwok AT, Lobo DN. Three centuries of appendicectomy. *World J Surg.* 2023;47:928-36.
7. McBurney C. Experience with early operative interference in cases of disease of the vermiform appendix. *Ny state med J.* 1889;50:676-84.
8. McBurney C. IV. The incision made in the abdominal wall in cases of appendicitis, with a description of a new method of operating. *Ann Surg.* 1894;20:38-43.
9. Bailey H. The Ochsner-Sherren (delayed) treatment of acute appendicitis: indications and technique. *Br Med J.* 1930;1:140-3.
10. Coldrey E. Treatment of Acute Appendicitis. *Br Med J.* 1956;2:1458-61.
11. Semm K. Endoscopic appendectomy. *Endoscopy.* 1983;15:59-64.
12. Reddy N, Rao P. Per oral transgastric endoscopic appendectomy in human. Rad prezentiran na: 45th Annual Conference of the Society of Gastrointestinal Endoscopy of India; 2004 Feb 28-29; Jaipur, India.
13. Jeong CW, Lee S, Oh JJ, Lee SB, Jeong SJ, Hong SK i sur. Pure transvesical NOTES appendectomy using a 5-mm rigid laparoscope: a feasibility and survival study with porcine models. *Surg Endosc.* 2013;27:2992-9.
14. Switzer NJ, Gill RS, Karmali S. The evolution of the appendectomy: from open to laparoscopic to single incision. *Scientifica (Cairo)*. 2012;2012:895469.
15. Chen T, Xu A, Lian J, Chu Y, Zhang H, Xu M. Transcolonic endoscopic appendectomy: a novel natural orifice transluminal endoscopic surgery (NOTES) technique for the sessile serrated lesions involving the appendiceal orifice. *Gut.* 2021;70:1812-4.
16. Palanivelu C, Rajan PS, Rangarajan M, Parthasarathi R, Senthilnathan P, Prasad M. Transvaginal endoscopic appendectomy in humans: a unique approach to NOTES--world's first report. *Surg Endosc.* 2008;22:1343-7.

17. Johnson LR. Encyclopedia of gastroenterology. 2. izdanje. Boston: Academic Press; 2004. 409 str.
18. Kacprzyk A, Droś J, Stefura T, Krzysztofik M, Jasińska K, Pędziwiatr M i sur. Variations and morphometric features of the vermiform appendix: A systematic review and meta-analysis of 114,080 subjects with clinical implications. Clin Anat. 2020;33:85-98.
19. Malas MA, Sulak O, Gokcimen A, Sari A. Development of the vermiform appendix during the fetal period. Surg Radiol Anat. 2004;26:202-7.
20. Delaney CP, Netter FH. Netter's surgical anatomy and approaches. 2. izdanje. Philadelphia: Elsevier; 2021. 746 str.
21. Wyllie R, Hyams JS, Kay M. Pediatric gastrointestinal and liver disease. 6. izdanje. Philadelphia: Elsevier; 2021. 541 str.
22. Yeo CJ. Shackelford's surgery of the alimentary tract. 8. izdanje. Philadelphia: Elsevier; 2019. 1951., 1162-1169 str.
23. Kooij IA, Sahami S, Meijer SL, Buskens CJ, Te Velde AA. The immunology of the vermiform appendix: a review of the literature. Clin Exp Immunol. 2016;186:1-9.
24. Constantin M, Petrescu L, Mătanie C, Vrancianu CO, Niculescu AG, Andronic O i sur. The Vermiform Appendix and Its Pathologies. Cancers (Basel). 2023;15:3872.
25. Smith HF. A review of the function and evolution of the cecal appendix. Anat Rec (Hoboken). 2023;306:972-82.
26. Barlow A, Muhleman M, Gielecki J, Matusz P, Tubbs RS, Loukas M. The vermiform appendix: a review. Clin Anat. 2013;26:833-42.
27. Randal Bollinger R, Barbas AS, Bush EL, Lin SS, Parker W. Biofilms in the large bowel suggest an apparent function of the human vermiform appendix. J Theor Biol. 2007;249:826-31.
28. Im GY, Modayil RJ, Lin CT, Geier SJ, Katz DS, Feuerman M i sur. The appendix may protect against clostridium difficile recurrence. Clin Gastroenterol Hepatol. 2011;9:1072-7.
29. Sanders NL, Bollinger RR, Lee R, Thomas S, Parker W. Appendectomy and Clostridium difficile colitis: relationships revealed by clinical observations and immunology. World J Gastroenterol. 2013;19:5607-14.
30. Shaikh DH, Patel H, Munshi R, Sun H, Mehershahi S, Baiomi A i sur. Patients with Clostridium difficile infection and prior appendectomy may be prone to worse outcomes. World J Gastrointest Surg. 2021;13:1436-47.

31. Arjomand Fard N, Armstrong H, Perry T, Wine E. Appendix and ulcerative colitis: a key to explaining the pathogenesis and directing novel therapies? *Inflamm Bowel Dis.* 2023;29:151-60.
32. Sahami S, Kooij IA, Meijer SL, Van den Brink GR, Buskens CJ, Te Velde AA. The link between the appendix and ulcerative colitis: clinical relevance and potential immunological mechanisms. *Am J Gastroenterol.* 2016;111:163-9.
33. Andersson RE, Olaison G, Tysk C, Ekbom A. Appendectomy and protection against ulcerative colitis. *N Engl J Med.* 2001;344:808-14.
34. Myrelid P, Landerholm K, Nordenvall C, Pinkney TD, Andersson RE. Appendectomy and the risk of colectomy in ulcerative colitis: a national cohort study. *Am J Gastroenterol.* 2017;112:1311-9.
35. Frisch M, Gridley G. Appendectomy in adulthood and the risk of inflammatory bowel diseases. *Scand J Gastroenterol.* 2002;37:1175-7.
36. Andersson RE, Olaison G, Tysk C, Ekbom A. Appendectomy is followed by increased risk of Crohn's disease. *Gastroenterology.* 2003;124:40-6.
37. Kaplan GG, Jackson T, Sands BE, Frisch M, Andersson RE, Korzenik J. The risk of developing Crohn's disease after an appendectomy: a meta-analysis. *Am J Gastroenterol.* 2008;103:2925-31.
38. Kaplan GG, Pedersen BV, Andersson RE, Sands BE, Korzenik J, Frisch M. The risk of developing Crohn's disease after an appendectomy: a population-based cohort study in Sweden and Denmark. *Gut.* 2007;56:1387-92.
39. Chung WS, Chung S, Hsu CY, Lin CL. Risk of inflammatory bowel disease following appendectomy in adulthood. *Front Med (Lausanne).* 2021;8:661752.
40. Fantodji C, Jantchou P, Parent ME, Rousseau MC. Appendectomy and risk for inflammatory bowel disease: effect of age and time post appendectomy - a cohort study. *BMJ Open Gastroenterol.* 2022;9:e000925.
41. Visser E, Heuthorst L, Pathmakanthan SP, Wildenberg ME, D'Haens GR, Dijkgraaf MGW i sur. P955 The effect of appendicectomy on the clinical course of Ulcerative Colitis: preliminary results. *J Crohns Colitis.* 2024;18:i1734-i.
42. Nageswaran H, Khan U, Hill F, Maw A. Appendiceal duplication: a comprehensive review of published cases and clinical recommendations. *World J Surg.* 2018;42:574-81.
43. Elsherbini AT, Atta MA, Elshahat S, Emara MH. Atypical duplex appendix arising from the ascending colon: a case report. *J Med Case Rep.* 2024;18:188.

-
44. Ngulube A, Ntoto COM, Matsika D, Ndebele W, Dube NS, Gapu P. A case report of complete appendiceal duplication on the normal site of a single caecum: a new variant? *Int J Surg Case Rep.* 2020;74:168-72.
 45. Ghali MS, Saleem N, Khalaf MH, Alkubaisi IK, Ali AHJ, Al Obahi M i sur. Congenital absence of appendix: a rare condition that could result in severe complications-a review of literature. *J Surg Case Rep.* 2023;2023:rjad661.
 46. Uriev L, Maslovsky I, Mnouskin Y, Ben-Dor D. Triple-barreled type of appendiceal triplication. *Ann Diag Pathol.* 2006;10:160-1.
 47. Collins DC. Agensis of the vermiform appendix. *Am J Surg.* 1951;82:689-96.
 48. Deshmukh S, Verde F, Johnson PT, Fishman EK, Macura KJ. Anatomical variants and pathologies of the vermiform appendix. *Emerg Radiol.* 2014;21:543-52.
 49. Hammood ZD, Salih AM, Mahal LA, Yas YT, Ghaleb HA, Kakamad FH. Agensis of vermiform appendix; a case report with literature review. *Int J Surg Case Rep.* 2021;87:106364.
 50. Lipton S, Estrin J, Glasser I. Diverticular disease of the appendix. *Surg Gynecol Obstet.* 1989;168:13-6.
 51. Altieri ML, Piozzi GN, Salvatori P, Mirra M, Piccolo G, Olivari N. Appendiceal diverticulitis, a rare relevant pathology: Presentation of a case report and review of the literature. *Int J Surg Case Rep.* 2017;33:31-4.
 52. Allahqoli L, Mazidimoradi A, Momenimovahed Z, Günther V, Ackermann J, Salehiniya H i sur. Appendiceal endometriosis: a comprehensive review of the literature. *Diagnostics (Basel).* 2023;13:1827.
 53. Patil DT. Non-Neoplastic and Neoplastic Disorders of the Appendix. U: Srivastava A, Allende DS, editors. *Gastrointestinal and Liver Pathology.* 3. izdanje. Philadelphia: Elsevier; 2024. 211-213 str.
 54. Hussein MRA, Al Bshabshe A, Elhakeem AA, Elsamman MK. Neurogenic appendicitis: a reappraisal of the clinicopathological features and pathogenesis. *Diagnostics (Basel).* 2022;12:1386.
 55. Misdraji J, Graeme-Cook FM. Miscellaneous conditions of the appendix. *Semin Diagn Pathol.* 2004;21:151-63.
 56. Endo K, Sato M, Saga K, Higashio A, Yuba Y, Morotomi Y. Torsion of vermiform appendix: case report and review of the literature. *Surg Case Rep.* 2020;6:6.
 57. Chaar CI, Wexelman B, Zuckerman K, Longo W. Intussusception of the appendix: comprehensive review of the literature. *Am J Surg.* 2009;198:122-8.
-

-
58. Lokuhetty D, White VA, Watanabe R, Cree IA. Digestive system tumours. 5. izdanje. Lyon: International Agency for Research on Cancer; 2019. 145 str.
 59. Amin MB, Edge SB, American Joint Committee on Cancer. AJCC cancer staging manual. 8. Izdanje. Cham: Springer; 2017. 239 str.
 60. Chauhan A, Chan K, Halfdanarson TR, Bellizzi AM, Rindi G, O'Toole D i sur. Critical updates in neuroendocrine tumors: Version 9 American Joint Committee on Cancer staging system for gastroenteropancreatic neuroendocrine tumors. *CA Cancer J Clin.* 2024;74:359-67.
 61. Mohamed A, Wu S, Hamid M, Mahipal A, Cjakrabarti S, Bajor D i sur. Management of appendix neuroendocrine neoplasms: insights on the current guidelines. *Cancers (Basel).* 2022;15:295.
 62. Di Saverio S, Podda M, De Simone B, Ceresoli M, Augustin G, Gori A i sur. Diagnosis and treatment of acute appendicitis: 2020 update of the WSES Jerusalem guidelines. *World J Emerg Surg.* 2020;15:27.
 63. Wickramasinghe DP, Xavier C, Samarasekera DN. The worldwide epidemiology of acute appendicitis: an analysis of the global health data exchange dataset. *World J Surg.* 2021;45:1999-2008.
 64. Global Burden of Disease Collaborator Group. Trends and levels of the global, regional, and national burden of appendicitis between 1990 and 2021: findings from the Global Burden of Disease Study 2021. *Lancet Gastroenterol Hepatol.* 2024;9:825-58.
 65. Brunicaudi FC, Andersen DK, Billiar TR, Dunn DL, Hunter JG, Kao LS i sur. Schwartz's principles of surgery. 11. izdanje. New York: McGraw-Hill; 2019. 1131-6 str.
 66. Coccolini F, Catena F. Textbook of emergency general surgery 1. izdanje. Cham: Springer; 2023. 984-5 str.
 67. Zinner MJ, Ashley SW, Joe Hines O. Maingot's abdominal operations. 13. izdanje. New York: McGraw-Hill Education; 2019. 737-744 str.
 68. Aoun R, Akel R, Kassar S, Mouawad C, Dahboul H, Chakhtoura G i sur. Acute appendicitis caused by metastasis from other organs: a systematic review. *J Gastrointest Cancer.* 2024;55:143-52.
 69. Habachi G, Aziza B, Ben-Ammar S, Maherzi O, Houas Y, Kerkeni Y i sur. Acute peritonitis secondary to post-traumatic appendicitis: a case report and literature review. *World J Clin Cases.* 2023;11:6491-7.
 70. Lamps LW. Infectious causes of appendicitis. *Infect Dis Clin North Am.* 2010;24:995-1018, ix-x.
-

71. Andersson RE, Agiorgiti M, Bendtsen M. Spontaneous resolution of uncomplicated appendicitis may explain increase in proportion of complicated appendicitis during covid-19 pandemic: a systematic review and meta-analysis. *World J Surg.* 2023;47:1901-16.
72. Gorter RR, Eker HH, Gorter-Stam MA, Abis GS, Acharya A, Ankersmit M i sur. Diagnosis and management of acute appendicitis. EAES consensus development conference 2015. *Surg Endosc.* 2016;30:4668-90. doi: 10.1007/s00464-016-5245-7.
73. Correa J, Jimeno J, Vallverdu H, Bizzoca C, Collado-Roura F, Estalella L i sur. Correlation between intraoperative surgical diagnosis of complicated acute appendicitis and the pathology report: clinical implications. *Surg Infect (Larchmt).* 2015;16:41-4.
74. Moris D, Paulson EK, Pappas TN. Diagnosis and management of acute appendicitis in adults: a review. *JAMA.* 2021;326:2299-311.
75. Snyder MJ, Guthrie M, Cagle S. Acute appendicitis: efficient diagnosis and management. *Am Fam Physician.* 2018;98:25-33.
76. Petroianu A. Diagnosis of acute appendicitis. *Int J Surg.* 2012;10:115-9.
77. Augustin G. Acute abdomen during pregnancy. 3. izdanje. Cham: Springer; 2023. 377-80.
78. Petrone P, Brathwaite CEM. Acute care surgery in geriatric patients. 1. izdanje. Cham: Springer; 2023. 389-90 str.
79. Manatakis DK, Aheimastos V, Antonopoulou M-I, Agalianos C, Tsiaoussis J, Xynos E. Unfinished business: a systematic review of stump appendicitis. *World J Surg.* 2019;43:2756-61.
80. Dikicier E, Altintoprak F, Ozdemir K, Gundogdu K, Uzunoglu MY, Cakmak G i sur. Stump appendicitis: a retrospective review of 3130 consecutive appendectomy cases. *World J Emerg Surg.* 2018;13:22.
81. Hajibandeh S, Hajibandeh S, Marshall MJ, Smart NJ, Winyard PG, Hyde C i sur. Biomarkers for diagnosis of acute appendicitis in adults. *Cochrane Database Syst Rev.* 2021;2021:CD011592.
82. Hodge SVL, Mickiewicz B, Lau M, Jenne CN, Thompson GC. Novel molecular biomarkers and diagnosis of acute appendicitis in children. *Biomark Med.* 2021;15:1055-65.
83. Skjold-Ødegaard B, Søreide K. The diagnostic differentiation challenge in acute appendicitis: how to distinguish between uncomplicated and complicated appendicitis in adults. *Diagnostics (Basel).* 2022;12:1724.

84. Lamm R, Kumar SS, Collings AT, Haskins IN, Abou-Setta A, Narula N i sur. Diagnosis and treatment of appendicitis: systematic review and meta-analysis. *Surg Endosc.* 2023;37:8933-90.
85. Gan DEY, Nik Mahmood NRK, Chuah JA, Hayati F. Performance and diagnostic accuracy of scoring systems in adult patients with suspected appendicitis. *Langenbecks Arch Surg.* 2023;408:267.
86. Kaewlai R, Tongsai S, Teerasamit W, Wongsangchan D, Noppakunsomboon N, Khamman P i sur. Validation of scoring systems for the prediction of complicated appendicitis in adults using clinical and computed tomographic findings. *Insights Imaging.* 2023;14:191.
87. Kobayashi T, Hidaka E, Ando A, Koganezawa I, Nakagawa M, Yokozuka K i sur. Preoperative scoring system for prediction of extended resection during emergency surgery for acute appendicitis. *Langenbecks Arch Surg.* 2023;408:443.
88. Kobayashi T, Hidaka E, Koganezawa I, Nakagawa M, Yokozuka K, Ochiai S i sur. Prediction model for failure of nonoperative management of uncomplicated appendicitis in adults. *World J Surg.* 2021;45:3041-7.
89. Andersson M, Andersson RE. The appendicitis inflammatory response score: a tool for the diagnosis of acute appendicitis that outperforms the Alvarado score. *World J Surg.* 2008;32:1843-9.
90. Andersson M, Kolodziej B, Andersson RE. Validation of the appendicitis inflammatory response (AIR) score. *World J Surg.* 2021;45:2081-91.
91. Dahiya DS, Akram H, Goyal A, Khan AM, Shahnoor S, Hassan KM i sur. Controversies and future directions in management of acute appendicitis: an updated comprehensive review. *J Clin Med.* 2024;13:3034.
92. Xu H, Yang S, Xing J, Wang Y, Sun W, Rong L i sur. Comparison of the efficacy and safety of antibiotic treatment and appendectomy for acute uncomplicated appendicitis: a systematic review and meta-analysis. *BMC Surg.* 2023;23:208.
93. Dilek ON, Kar H, Acar T. Prophylactic appendectomy. In: Dilek ON, Uranues S, Latifi R, editors. *Prophylactic surgery.* Cham: Springer; 2021. 181-91 str.
94. Podda M, Gerardi C, Cillara N, Fearnhead N, Gomes CA, Birindelli A i sur. Antibiotic treatment and appendectomy for uncomplicated acute appendicitis in adults and children: a systematic review and meta-analysis. *Ann Surg.* 2019;270:1028-40.
95. Brucchi F, Bracchetti G, Fugazzola P, Viganò J, Filisetti C, Ansaloni L i sur. A meta-analysis and trial sequential analysis comparing nonoperative versus operative

- management for uncomplicated appendicitis: a focus on randomized controlled trials. *World J Emerg Surg.* 2024;19:2. doi: 10.1186/s13017-023-00531-6.
96. Doleman B, Fonnes S, Lund JN, Boyd-Carson H, Javanmard-Emamghissi H, Moug S i sur. Appendectomy versus antibiotic treatment for acute appendicitis. *Cochrane Database Syst Rev.* 2024;4:CD015038.
97. Park HC, Kim MJ, Lee BH. Randomized clinical trial of antibiotic therapy for uncomplicated appendicitis. *Br J Surg.* 2017;104:1785-90.
98. Salminen P, Sippola S, Haijanen J, Nordström P, Rantanen T, Rautio T i sur. Antibiotics versus placebo in adults with CT-confirmed uncomplicated acute appendicitis (APPAC III): randomized double-blind superiority trial. *Br J Surg.* 2022;109:503-9.
99. Iresjö B-M, Blomström S, Engström C, Johnsson E, Lundholm K. Acute appendicitis: a block-randomized study on active observation with or without antibiotic treatment. *Surgery.* 2024;175:929-35.
100. Weis JJ, Hamilton EC. Appendicitis. U: Renton D, Lim R, Gallo AS, Sinha P, urednici. *The SAGES Manual of Acute Care Surgery.* 1. izdanje. Cham: Springer International Publishing; 2020. 15-32 str.
101. Apeltgren KN. Laparoscopic appendectomy. U: Scott-Conner CEH, urednik. *The SAGES manual: fundamentals of laparoscopy, thoracoscopy, and GI endoscopy.* New York: Springer New York; 2006. 350-6 str.
102. Davis CR, Trevatt A, Dixit A, Datta V. Systematic review of clinical outcomes after prophylactic surgery. *Ann R Coll Surg Engl.* 2016;98:353-7.
103. Reyes DP, Carroll DJ, Walton ME, Antonsen EL, Kerstman EL. Probabilistic risk assessment of prophylactic surgery before extended-duration spaceflight. *Surg Innov.* 2021;28:573-81.
104. Theodosiou AA, Read RC. Artificial intelligence, machine learning and deep learning: potential resources for the infection clinician. *J Infect.* 2023;87:287-94.
105. Vijayalakshmi A. Lepakshi. Machine learning and deep learning based AI tools for development of diagnostic tools. U: Parihar A, Khan R, Kumar A, Kaushik AK, Gohel H, urednici. *Computational approaches for novel therapeutic and diagnostic designing to mitigate SARS-CoV-2 infection.* 1. izdanje. Cambridge (Massachusetts): Academic Press; 2022. 399–420 str.
106. Sarker IH. AI-based modeling: techniques, applications and research issues towards automation, intelligent and smart systems. *SN Comput Sci.* 2022;3:158.

-
107. An Q, Rahman S, Zhou J, Kang JJ. A comprehensive review on machine learning in healthcare industry: classification, restrictions, opportunities and challenges. *Sensors (Basel)*. 2023;23:4178.
108. Vrdoljak, J. (2023). *Modeli strojnog učenja za predviđanje metastaza u limfne čvorove u bolesnika s rakom dojke (Disertacija)*. Split: Sveučilište u Splitu, Medicinski fakultet. Dostupno na: <https://urn.nsk.hr/urn:nbn:hr:171:497333>.
109. Black JE, Kueper JK, Williamson TS. An introduction to machine learning for classification and prediction. *Fam Pract*. 2022;40:200-4.
110. Jovel J, Greiner R. An introduction to machine learning approaches for biomedical research. *Front med (Lausanne)*. 2021;8:771607.
111. Sutton R, Barto A. Reinforcement learning: an introduction. 2. izdanje. MIT Press; 2018 [citirano 13. rujna 2024.]. dostupno na: <http://incompleteideas.net/book/the-book-2nd.html>.
112. Al-Azzam N, Shatnawi I. Comparing supervised and semi-supervised machine learning models on diagnosing breast cancer. *Ann Med Surg (Lond)*. 2021;62:53-64.
113. Morris MX, Rajesh A, Asaad M, Hassan A, Saadoun R, Butler CE. Deep learning applications in surgery: current uses and future directions. *Am Surg*. 2023;89:36-42.
114. Choi RY, Coyner AS, Kalpathy-Cramer J, Chiang MF, Campbell JP. Introduction to machine learning, neural networks, and deep learning. *Transl Vis Sci Technol*. 2020;9:14.
115. Jurafsky D, Martin JH. Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition with language models [Internet]. 3. izdanje. 2024. Dostupno na: <https://web.stanford.edu/~jurafsky/slp3/>.
116. Blockeel H, Devos L, Frénay B, Nanfack G, Nijssen S. Decision trees: from efficient prediction to responsible AI. *Front Artif Intell*. 2023;6:1124553.
117. Moore A, Bell M. Xgboost, a novel explainable ai technique, in the prediction of myocardial infarction: a UK biobank cohort study. *Clin Med Insights Cardiol*. 2022;16:11795468221133611.
118. Matsuki K, Kuperman V, Van Dyke JA. The random forests statistical technique: an examination of its value for the study of reading. *Sci Stud Read*. 2016;20:20-33.
119. Bentéjac C, Csörgő A, Martínez-Muñoz G. A comparative analysis of gradient boosting algorithms. *Artif Intell Rev*. 2021;54:1937-67.
120. Maleki F, Muthukrishnan N, Ovens K, Reinhold C, Forghani R. Machine learning algorithm validation: from essentials to advanced applications and implications for regulatory certification and deployment. *Neuroimaging Clin N Am*. 2020;30:433-45.
-

-
- 121.Montesinos López OA, Montesinos López A, Crossa J. Overfitting, model tuning, and evaluation of prediction performance. Multivariate statistical machine learning methods for genomic prediction. Cham: Springer International Publishing; 2022. 109-39 str.
 - 122.Hicks SA, Strümke I, Thambawita V, Hammou M, Riegler MA, Halvorsen P i sur. On evaluation metrics for medical applications of artificial intelligence. *Sci Rep.* 2022;12:5979.
 - 123.Christen P, Hand DJ, Kirielle N. A review of the F-measure: its history, properties, criticism, and alternatives. *ACM Comput Surv.* 2023;56:73.
 - 124.Rozemberczki B, Watson L, Bayer P, Yang H-T, Kiss O, Nilsson S i sur. The Shapley value in machine learning. U: De Raedt L, urednik. Proceedings of the 31st International Joint Conference on Artificial Intelligence, IJCAI-ECAI 2022. Beč (Austrija): International Joint Conferences on Artificial Intelligence Organization; 2022. 5572–5579 str. Dostupno na: <https://doi.org/10.24963/ijcai.2022/778>:
 - 125.Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. U: Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS 2017). Red Hook: Curran Associates Inc.; 2017. 4768–4777 str. Dostupno na: <https://ui.adsabs.harvard.edu/abs/2017arXiv170507874L>.
 - 126.Habebh H, Gohel S. Machine learning in healthcare. *Curr Genomics.* 2021;22:291-300.
 - 127.Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: promises and perils. *Ann Surg.* 2018;268:70-6.
 - 128.Dayan D, Dvir N, Agbariya H, Nizri E. Implementation of artificial intelligence-based computer vision model in laparoscopic appendectomy: validation, reliability, and clinical correlation. *Surg Endosc.* 2024;38:3310-9.
 - 129.Nyangoh Timoh K, Huaulme A, Cleary K, Zaheer MA, Lavoué V, Donoho D i sur. A systematic review of annotation for surgical process model analysis in minimally invasive surgery based on video. *Surg Endosc.* 2023;37:4298-314.
 - 130.Ortenzi M, Rapoport Ferman J, Antolin A, Bar O, Zohar M, Perry O i sur. A novel high accuracy model for automatic surgical workflow recognition using artificial intelligence in laparoscopic totally extraperitoneal inguinal hernia repair (TEP). *Surg Endosc.* 2023;37:8818-28.
 - 131.Bianchi V, Giambusso M, De Iacob A, Chiarello MM, Brisinda G. Artificial intelligence in the diagnosis and treatment of acute appendicitis: a narrative review. *Updates Surg.* 2024;76:783-92.

-
132. Aydin E, Türkmen İ U, Namli G, Öztürk Ç, Esen AB, Eray YN i sur. A novel and simple machine learning algorithm for preoperative diagnosis of acute appendicitis in children. *Pediatr Surg Int.* 2020;36:735-42.
133. Mijwil MM, Aggarwal K. A diagnostic testing for people with appendicitis using machine learning techniques. *Multimed Tools Appl.* 2022;81:7011-23.
134. Akbulut S, Yagin FH, Cicek IB, Koc C, Colak C, Yilmaz S. Prediction of perforated and nonperforated acute appendicitis using machine learning-based explainable artificial intelligence. *Diagnostics (Basel).* 2023;13:1173.
135. Ghareeb WM, Emile SH, Elshobaky A. Artificial intelligence compared to alvarado scoring system alone or combined with ultrasound criteria in the diagnosis of acute appendicitis. *J Gastrointest Surg.* 2022;26:655-8.
136. Reismann J, Romualdi A, Kiss N, Minderjahn MI, Kallarackal J, Schad M i sur. Diagnosis and classification of pediatric acute appendicitis by artificial intelligence methods: an investigator-independent approach. *PLoS One.* 2019;14:e0222030.
137. Park SY, Kim SM. Acute appendicitis diagnosis using artificial neural networks. *Technol Health Care.* 2015;23 Suppl 2:S559-65.
138. Prabhudesai SG, Gould S, Rekhraj S, Tekkis PP, Glazer G, Ziprin P. Artificial neural networks: useful aid in diagnosing acute appendicitis. *World J Surg.* 2008;32:305-9; discussion 10-1.
139. Yoldaş Ö, Tez M, Karaca T. Artificial neural networks in the diagnosis of acute appendicitis. *Am J Emerg Med.* 2012;30:1245-7.
140. Sakai S, Kobayashi K, Toyabe S, Mandai N, Kanda T, Akazawa K. Comparison of the levels of accuracy of an artificial neural network model and a logistic regression model for the diagnosis of acute appendicitis. *J Med Syst.* 2007;31:357-64.
141. Akmeşe OF, Dogan G, Kor H, Erbay H, Demir E. The use of machine learning approaches for the diagnosis of acute appendicitis. *Emerg Med Int.* 2020;2020:7306435.
142. Phan-Mai TA, Thai TT, Mai TQ, Vu KA, Mai CC, Nguyen DA. Validity of machine learning in detecting complicated appendicitis in a resource-limited setting: findings from Vietnam. *Biomed Res Int.* 2023;2023:5013812.
143. Hsieh CH, Lu RH, Lee NH, Chiu WT, Hsu MH, Li YC. Novel solutions for an old disease: diagnosis of acute appendicitis with random forest, support vector machines, and artificial neural networks. *Surgery.* 2011;149:87-93.
-

-
144. Stiel C, Elrod J, Klinkle M, Herrmann J, Junge CM, Ghadban T i sur. The modified Heidelberg and the AI appendicitis score are superior to current scores in predicting appendicitis in children: a two-center cohort study. *Front Pediatr*. 2020;8:592892.
145. Rajpurkar P, Park A, Irvin J, Chute C, Bereket M, Mastrodicasa D i sur. AppendiXNet: deep learning for diagnosis of appendicitis from a small dataset of CT exams using video pretraining. *Sci Rep*. 2020;10:3958.
146. Reismann J, Kiss N, Reismann M. The application of artificial intelligence methods to gene expression data for differentiation of uncomplicated and complicated appendicitis in children and adolescents - a proof of concept study. *BMC Pediatr*. 2021;21:268.
147. Kang CB, Li XW, Hou SY, Chi XQ, Shan HF, Zhang QJ i sur. Preoperatively predicting the pathological types of acute appendicitis using machine learning based on peripheral blood biomarkers and clinical features: a retrospective study. *Ann Transl Med*. 2021;9:835.
148. Marcinkevics R, Reis Wolfertstetter P, Wellmann S, Knorr C, Vogt JE. Using machine learning to predict the diagnosis, management and severity of pediatric appendicitis. *Front Pediatr*. 2021;9:662183.
149. Marcinkevics R, Reis Wolfertstetter P, Klimiene U, Chin-Cheong K, Paschke A, Zerres J i sur. Interpretable and intervenable ultrasonography-based machine learning models for pediatric appendicitis. *Med Image Anal*. 2024;91:103042.
150. Liang D, Fan Y, Zeng Y, Zhou H, Zhou H, Li G i sur. Development and validation of a deep learning and radiomics combined model for differentiating complicated from uncomplicated acute appendicitis. *Acad Radiol*. 2024;31:1344-54.
151. Maleš I, Kumrić M, Huić Maleš A, Cvitković I, Šantić R, Pogorelić Z i sur. A systematic integration of artificial intelligence models in appendicitis management: a comprehensive review. *Diagnostics (Basel)*. 2025;15:866.
152. Schipper A, Belgers P, O'Connor R, Jie KE, Dooijes R, Bosma JS i sur. Machine-learning based prediction of appendicitis for patients presenting with acute abdominal pain at the emergency department. *World J Emerg Surg*. 2024;19:40.
153. Navaei M, Doogchi Z, Gholami F, Tavakoli MK. Leveraging machine learning for pediatric appendicitis diagnosis: a retrospective study integrating clinical, laboratory, and imaging data. *Health Sci Rep*. 2025;8:e70756.
154. Henriksen SR, Christophersen C, Rosenberg J, Fonnes S. Varying negative appendectomy rates after laparoscopic appendectomy: a systematic review and meta-analysis. *Langenbecks Arch Surg*. 2023;408:205.
-

- 155.Kinesya E, Cintya EP, Dorothy MJ, Ennaidi NN, Rusti HF, Mannagalli Y i sur. Diagnostic accuracy of Alvarado score components in patients with appendicitis: systematic review and meta-analysis approach. *Health Sci Rev.* 2022;2:100018.
- 156.Kendall J, Gaspar G, Berger D, Levman J. Machine learning and feature selection in pediatric appendicitis. *Tomography.* 2025;11:90.
- 157.Chadaga K, Khanna V, Prabhu S, Sampathila N, Chadaga R, Umakanth S i sur. An interpretable and transparent machine learning framework for appendicitis detection in pediatric patients. *Sci Rep.* 2024;14:24454.
- 158.Alkhawaldeh IM, Albalkhi I, Naswhan AJ. Challenges and limitations of synthetic minority oversampling techniques in machine learning. *World J Methodol.* 2023;13:373-8.
- 159.Erman A, Ferreira J, Ashour WA, Guadagno E, St-Louis E, Emil S i sur. Machine-learning-assisted preoperative prediction of pediatric appendicitis severity. *J Pediatr Surg.* 2025;60:162151.
- 160.Yeh DD, Eid AI, Young KA, Wild J, Kaafarani HMA, Ray-Zack M i sur. Multicenter study of the treatment of appendicitis in america: acute, perforated, and gangrenous (MUSTANG), an EAST multicenter study. *Ann Surg.* 2021;273:548-56.
- 161.Minneci PC, Hade EM, Lawrence AE, Sebastião YV, Saito JM, Mak GZ i sur. Association of nonoperative management using antibiotic therapy vs laparoscopic appendectomy with treatment success and disability days in children with uncomplicated appendicitis. *Jama.* 2020;324:581-93.
- 162.Li J. Revisiting delayed appendectomy in patients with acute appendicitis. *World J Clin Cases.* 2021;9:5372-90.
- 163.Jumah S, Wester T. Non-operative management of acute appendicitis in children. *Pediatr Surg Int.* 2022;39:11.
- 164.Wei W, Tongping S, Jiaming W. Construction of a clinical prediction model for complicated appendicitis based on machine learning techniques. *Sci Rep.* 2024;14:16473.
- 165.Chen S, Xia J, Xu B, Huang Y, Teng M, Pan J. Risk prediction and effect evaluation of complicated appendicitis based on XGBoost modeling. *BMC Gastroenterol.* 2025;25:295.
- 166.Gollapalli M, Rahman A, Kudos SA, Foula MS, Alkhalifa AM, Albisher HM i sur. Appendicitis diagnosis: ensemble machine learning and explainable artificial intelligence-based comprehensive approach. *Big Data Cogn Comp.* 2024;8:108.

-
- 167.Xia J, Wang Z, Yang D, Li R, Liang G, Chen H i sur. Performance optimization of support vector machine with oppositional grasshopper optimization for acute appendicitis diagnosis. *Comput Biol Med.* 2022;143:105206.
- 168.Aydın E, Sarnıç TE, Türkmen İU, Khanmammadova N, Ateş U, Öztan MO i sur. Diagnostic accuracy of a machine learning-derived appendicitis score in children: a multicenter validation study. *Children.* 2025;12:937.
- 169.Kucukakcali Z, Akbulut S, Colak C. Evaluating ensemble-based machine learning models for diagnosing pediatric acute appendicitis: insights from a retrospective observational study. *J Clin Med.* 2025;14:4264.
- 170.Andersson RE, Stark J. Diagnostic value of the appendicitis inflammatory response (AIR) score. A systematic review and meta-analysis. *World J Emerg Surg.* 2025;20:12.
- 171.Kucukakcali Z, Cicek IB. Comparative diagnostic performance of RIPASA, ALVARADO, and AIR scoring systems in acute appendicitis: a retrospective evaluation using traditional metrics and XGBoost algorithm. *Int J Med All Body Health Res.* 2025;6:71-6.
- 172.Marcinkevics R. (2020) Pediatric Appendicitis Prediction Tool. Dostupno na: <http://papt.inf.ethz.ch/>.
- 173.Marcinkevics R, Sokol K, Paulraj A, Hilbert MA, Rimili V, Wellmann S i sur. External validation of predictive models for diagnosis, management and severity of pediatric appendicitis. *Front Pediatr.* 2025;13:1587488.
- 174.Jukić M, Nizeteo P, Matas J, Pogorelić Z. Trends and predictors of pediatric negative appendectomy rates: a single-centre retrospective study. *Children (Basel).* 2023;10:887.
- 175.Li Y, Yang Y, Song P, Duan L, Ren R. An improved SMOTE algorithm for enhanced imbalanced data classification by expanding sample generation space. *Sci Rep.* 2025;15:23521.
- 176.Jiang J, Zhang C, Ke L, Hayes N, Zhu Y, Qiu H i sur. A review of machine learning methods for imbalanced data challenges in chemistry. *Chem Sci.* 2025;16:7637-58.

OPĆI PODACI:

Ime i prezime: Ivan Maleš

Datum rođenja: 20. travnja 1994.

OBRAZOVANJE

2021. – polaznik doktorskog studija „Klinička medicina utemeljena na dokazima“, Medicinski fakultet u Splitu

2020. doktor medicine, Medicinski fakultet u Split

RADNO ISKUSTVO

2021. – danas, doktor medicine specijalizant abdominalne kirurgije, KBC Split

2023. – danas, vanjski suradnik, Katedra za kirurgiju, Medicinski fakultet Sveučilišta u Splitu

09/2020. – 03/2021. Rad pod nadzorom, Ordinacija obiteljske medicine dr. Diana Petričević

2018. – 2020. Demonstrator na Katedri za kliničke vještine

NASTAVNA AKTIVNOST

2023. – danas, Kirurgija (studij Medicina, studij Dentalna medicina, studij Medicina na engleskom jeziku), Medicinski fakultet, Sveučilište u Splitu

OSTALO

2025. – danas, istraživač, Laboratorij za umjetnu inteligenciju i podatkovnu znanost u biomedicini, Medicinski fakultet Sveučilišta u Splitu

2025. – danas, član, Povjerenstvo za primjenu UI tehnologije, Medicinski fakultet Sveučilišta u Splitu

NAGRADE I PRIZNANJA

2021. Rektorova nagrada za posebna postignuća

STRANI JEZICI

Engleski jezik (napredno), švedski jezik (osnovno)

KONGRESNA PRIOPĆENJA

1. Barcot O, Quien D, Boric M, Tavra A, **Males I**, Pavic P, Perko MA. Kronični presakralni sinus nakon niske resekcije rektuma: pregled literature i prikaz kliničkih slučajeva. *Acta Chir Croat.* 2025;22:14. Knjiga sažetaka. Izloženo na: 16. kongres Hrvatskog društva za digestivnu kirurgiju s međunarodnim sudjelovanjem; 2025; Opatija, Hrvatska.
2. **Males I**, Kumric M, Vrdoljak J, Perko MA, Bozic J. The role of artificial intelligence in the diagnosis and treatment of acute appendicitis. *Acta Chir Croat.* 2024;21:14. Knjiga sažetaka. Izloženo na: 16. kongres Hrvatskog društva za endoskopsku kirurgiju s međunarodnim sudjelovanjem; 2024; Dubrovnik, Hrvatska.
3. Majic A, **Males I**, Perko Z. Laparoscopic total gastrectomy for gastric cancer, side to side technique for reconstruction: a case report. *Acta Chir Croat.* 2024;21:49. Knjiga sažetaka. Izloženo na: 16. kongres Hrvatskog društva za endoskopsku kirurgiju s međunarodnim sudjelovanjem; 2024; Dubrovnik, Hrvatska.
4. Bucan Nenadic D, Selak M, Perica K, Kolak E, **Males I**, Perko Z. Nutritional intake of hospitalized patients with fracture: nutritionDay 2022 analysis. Poster prezentiran na: Clinical Nutrition ESPEN; 11–14.11.2023; Lyon, Francuska. *Clin Nutr ESPEN.* 2023;58:573.
5. **Males I**, Selak M, Bucan Nenadic D, Kolak E, Perica K, Perko Z. Incidence of perioperative malnutrition in elderly surgical patients. Poster prezentiran na: Clinical Nutrition ESPEN; 11–14.11.2023; Lyon, Francuska. *Clin Nutr ESPEN.* 2023;58:573.

PUBLIKACIJE

1. Bozic J, Kumric M, Ticinovic Kurir T, **Males I**, Borovac J, Martinovic D, Vilovic M. Role of adropin in cardiometabolic disorders: from pathophysiological mechanisms to therapeutic target. *Biomedicines.* 2021;9:1407.
2. Vrdoljak J, Kumric M, Ticinovic Kurir T, **Males I**, Martinovic D, Vilovic M, Bozic J. Effects of wine components in inflammatory bowel diseases. *Molecules.* 2021;26:5891.
3. Martinovic D, Tokic D, Martinovic L, Kumric M, Vilovic M, Rusic D, Vrdoljak J, **Males I**, Ticinovic Kurir T, Lupi-Ferandin S, Bozic J. Adherence to the mediterranean diet and its association with the level of physical activity in fitness center users: Croatian-based study. *Nutrients.* 2021;13:4038.
4. Tonkic A, Kumric M, Akrapovic Olic I, Rusic D, Marin Zivkovic P, Supe Domic D, Sundov Z, **Males I**, Bozic J. Growth differentiation factor-15 serum concentrations reflect

- disease severity and anemia in patients with inflammatory bowel disease. *World J Gastroenterol*. 2024;30:1899-910.
5. **Males I**, Boban Z, Kumric M, Vrdoljak J, Berkovic K, Pogorelic Z, Bozic J. Applying an explainable machine learning model might reduce the number of negative appendectomies in pediatric patients with a high probability of acute appendicitis. *Sci Rep*. 2024;14:12772.
 6. Vrdoljak J, Boban Z, **Males I**, Skrabic R, Kumric M, Ottosen A, Clemencau A, Bozic J, Völker S. Evaluating large language and large reasoning models as decision support tools in emergency internal medicine. *Comput Biol Med*. 2025;192:11035.
 7. **Males I**, Kumric M, Huic Males A, Cvitkovic I, Santic R, Pogorelic Z, Bozic J. A systematic integration of artificial intelligence models in appendicitis management: a comprehensive review. *Diagnostics*. 2025;15:866.
 8. Pogorelic Z., Jukic M, Zuvela T, Milunovic KP, **Males I**, Lovrinevic I i sur. Salivary biomarkers in pediatric acute appendicitis: current evidence and future directions. *Children*. 2025;12:1342.