

UNIVERZITET U BEOGRADU

Veličko V. Vraneš

OPTIMIZACIJA RAČUNARSKE ANALIZE MORFOLOGIJE
EPITELNIH MALIGNIH PLAŽA TUMORA DOJKE U
CILJU POBOLJŠANJA PROGNOZE ISHODA BOLESTI

doktorska disertacija

Beograd, 2022

UNIVERSITY OF BELGRADE

Veličko V. Vraneš

OPTIMIZATION OF COMPUTATIONAL ANALYSIS OF
MORPHOLOGY OF MALIGNANT EPITHELIAL NESTS
OF BREAST TUMORS IN ORDER TO IMPROVE
DISEASE PROGNOSIS

Doctoral Dissertation

Belgrade, 2022

Podaci o mentorima i članovima komisije

PODACI O MENTORIMA:

Prof. dr Nebojša Milošević, mentor

redovni profesor, Univerzitet u Beogradu, Medicinski fakultet

dr Marko Radulović, mentor

naučni savetnik, Institut za onkologiju i radiologiju Srbije, Beograd

PODACI O ČLANOVIMA KOMISIJE:

Prof. dr Zoran Nikolić, Univerzitet u Beograd, Fizički fakultet

Prof. dr Dejan Žikić, Univerzitet u Beogradu, Medicinski fakultet

dr Ana Damjanović Veličković, naučni saradnik, Institut za onkologiju i radiologiju Srbije, Beograd

Datum odbrane: _____

OPTIMIZACIJA RAČUNARSKE ANALIZE MORFOLOGIJE EPITELNIH MALIGNIH PLAŽA TUMORA DOJKE U CILJU POBOLJŠANJA PROGNOZE ISHODA BOLESTI

Sažetak: Napredak u preživljavanju i poboljšanju kvaliteta života pacijenata sa rakom dojke se može ostvariti personalizacijom terapije koja trenutno nije u punoj meri izvodljiva usled nedovoljne pouzdanosti prognostičkih pokazatelja. Ova teza je zato imala cilj da identifikuje izvore prognostičkih informacija i poboljša njihove prognostičke performanse na osnovu morfometrijskih osobina epitelnih malignih plaža, tj. malignih klastera u tumoru dojke i njihove ekspresije citokeratina.

Obuhvaćeno je 102 pacijentkinje bez sistemske terapije koja bi uticala na pojavu metastaza. Analizirane su 532 slike histopatoloških preparata tumora bojnih koktelom AE1/AE3 pan-citokeratinskih antitela čime je omogućena vizualizacija epitelnih ćelija koje su kod karcinoma dojke najčešće maligno transformisane. Da bi se ispitaio uticaj ekspresije citokeratina na prognostički učinak, izvršena je segmentacija slika na sedam nepreklapajućih opsega intenziteta piksela. Identifikaciji morfometrijskog profila (oblik, veličina, broj) malignih plaža sa najvećim prognostičkim značajem pristupilo se kombinacijom binarizacije slika sive skale sa šest pragova, pet filtera cirkularnosti i četiri filtera veličine. Tako je dobijeno 120 setova slika sa malignim plažama razdvojenim u odnosu na njihovu morfologiju. Prognostički učinak izračunatih parametara je utvrđen ROC analizom, a njihov značaj i nezavisnost su ispitani Cox regresionim modelom, koristeći pojavu udaljenih metastaza kao ishod bolesti. Značaj pojave metastaza je u tome što su metastaze, a ne primarni tumor, uzrok smrtnosti kod raka dojke.

Rezultati su pokazali koncentraciju prognostičkog značaja u epitelnim malignim plažama sa povećanom cirkularnošću i manjom veličinom, koje eksprimiraju citokeratine u srednjem i niskom intenzitetu. Pri tome su plaže sa najboljim prognostičkim performansama (AUC = 0.82) identifikovane kao pojedinačne epitelne ćelije rasute po stromi tumora. Njihov broj je bio najefikasniji i nezavisan prognostički parametar. Izuzetan prognostički značaj ovih ćelija je baziran i na činjenici da se one mogu identifikovati, kvantifikovati i analizirati ne samo računarskom analizom već i vizuelnim mikroskopskim pregledom.

Na kraju, analizom prognostičkog značaja morfometrije malignih tumorskih plaža ostvarena je prognostička tačnost ishoda bolesti znatno viša od tačnosti standardnih kliničkih parametara. Postignuto poboljšanje prognostičke efikasnosti upućuje na potencijalnu kliničku primenu dobijenih rezultata na osnovu pružanja objektivne podrške u personalizaciji terapije. Pouzdana identifikacija pacijenata niskog rizika od metastaza bi tako omogućila davanje manje intenzivnog tretmana i poboljšanje kvaliteta života, a kod onih sa visokim rizikom, davanje intenzivnijih terapija i mogućim produženjem preživljavanja.

Ključne reči: rak dojke, metastaze, pan-citokeratin, maligne epitelne plaže, individualne ćelije, analiza slike, prognoza, teksturalna i morfometrijska analiza, ROC-AUC analiza,

Naučna oblast: Biofizika. **Uža naučna oblast:** Digitalna histopatologija.

OPTIMIZATION OF COMPUTATIONAL ANALYSIS OF MORPHOLOGY OF MALIGNANT EPITHELIAL NESTS OF BREAST TUMORS IN ORDER TO IMPROVE DISEASE PROGNOSIS

Summary: Progress in the survival and life quality improvement of breast cancer patients could be achieved by personalized treatment, which is currently not fully achievable due to insufficient reliability of prognostic factors. This thesis aimed to identify sources of prognostic information and improve their prognostic performance based on the morphometric properties of epithelial malignant nests (i.e., malignant clusters) in breast tumor and their cytokeratin expression.

102 patients without systemic therapy that could affect the metastases occurrence were studied. 532 images of tumors tissue samples stained by the AE1/AE3 pan-cytokeratin antibody cocktail were analyzed, which enabled the visualization of epithelial cells that are most often malignantly transformed in breast cancer. To examine the effect of cytokeratin expression on prognostic efficiency, image segmentation was performed on seven non-overlapping pixel intensity ranges. The identification of the morphometric profile (shape, size, number) of malignant nests with the highest prognostic significance was accomplished applying six binarization thresholds in combination with five circularity filters and four size filters. Therefore, 120 sets of images were obtained with malignant nests separated based on their morphology. The prognostic performance of the extracted features was determined by ROC analysis, and their significance and independence were examined by Cox regression model, using distant metastasis occurrence as an endpoint event. The significance of metastasis occurrence is that metastasis, and not the primary tumor, is the cause of breast cancer mortality.

The results showed that malignant epithelial nests with greater circularity, smaller size and with medium and low intensity of cytokeratin expression, are the most important for prognosis. The nests with the best prognostic performance (AUC = 0.82) were identified as individual epithelial cells scattered in the tumor

stroma, and their count was the most efficient and independent prognostic factor. The exceptional prognostic significance of these cells is also based on the fact that they can be identified, quantified and studied either by computational analysis or visual microscopy.

In conclusion, the analysis of the prognostic importance of the malignant tumor nests morphometry yielded a prognostic accuracy that significantly exceeds the accuracy of existing standard clinical parameters. The achieved prognostic efficacy enables a clinical application of the obtained results by allowing greater objectivity in therapy personalization. Reliable identification of low-risk breast cancer patients could allow the prescription of less intense treatment and an improvement in quality of life, and for high-risk patients, more intense therapy and a possible increase in patient survival.

Key words: breast cancer, metastasis, pan-cytokeratin, malignant epithelial nests, individual cells, image analysis, prognosis, textural and morphometric analysis, ROC-AUC analysis,

Scientific field: Biophysics. **Scientific subfield:** Digital histopathology.

Sadržaj

1. UVOD.....	1
1.1. Prognostički faktori ishoda primarnog kancera dojke	4
1.1.1. Heterogenost tumora dojke i standardno korišćeni prognostički faktori	5
1.2. Računarska analiza digitalnih slika histopatoloških uzoraka raka dojke	9
1.2.1. Digitalna patologija i računarski prognostički faktori	9
2. HIPOTEZE ISTRAŽIVANJA.....	13
3. CILJEVI ISTRAŽIVANJA	14
4. MATERIJALI I METODE.....	16
4.1. Uzorak pacijenata	16
4.1.1. Testiranje HER2 amplifikacije	16
4.2. Protokol rada u računarskoj analizi digitalnih slika	17
4.3. Prikupljanje slika i selekcija regiona od interesa	18
4.3.1. Histopatološka priprema i bojenje uzoraka	18
4.3.2. Skeniranje uzoraka i izdvajanje reprezentativnih segmenata	19
4.4. Priprema slika za kvantitativnu analizu	19
4.4.1. Transformacija originalnih slika u slike sive skale i podela histograma	20
4.4.2. Binarizacija slika	21
4.5. Analiza slika i računanje parametara	24
4.5.1. Analiza teksture digitalnih slika – statistički prilaz	24
4.5.2. Računanje teksturalnih parametara	32
4.5.3. Savremena morfometrijska analiza histopatoloških slika	33
4.5.4. Računanje broja i morfometrijskih parametara epitelnih malignih plaža	37
4.6. Procena učinka prognostičkih faktora u medicini	38
4.6.1. ROC analiza u medicinskim istraživanjima	41
4.6.2. Cox-ov regresioni model proporcionalnog rizika	47
4.7. Obrada rezultata	51
5. REZULTATI	52
5.1. Kliničko-patološke karakteristike uzorka pacijentkinja	52
5.2. Prognostički učinak starosti pacijentkinja i kliničko-patoloških parametara	54
5.3. Raspodela intenziteta pan-citokeratinskog imunobojenja	55
5.3.1. Specifično i nespecifično pan-citokeratinsko imunobojenje	57

5.4. Prognoštički učinak teksturalnih parametra	59
5.4.1. Prognoštički učinak teksturalnih parametara u intervalima specifičnog i nespecifičnog pan-citokeratinskog imunobojenja	64
5.4.2. Poboľšanje prognoštičkog učinka teksturalnih parametara usled podele histograma i procena prognoštički najefikasnijeg intervala intenziteta imunobojenja.....	65
5.5. Prognoštički učinak morfometrijskih parametara epitelnih malignih plaža.....	67
5.5.1. Prognoštički učinak morfometrijskih parametara u funkciji praga binarizacije.....	67
5.5.2. Poboľšanje prognoštičkog učinka morfometrijskih parametara usled optimizacije praga binarizacije i procena prognoštički najefikasnijeg praga binarizacije.....	72
5.5.3. Prognoštički učinak morfometrijskih parametara pri optimalnoj kombinaciji vrednosti praga binarizacije i filtera veličine i cirkularnosti	74
5.5.4. Identifikacija prognoštički najznačajnijih čestica	77
5.6. Multivarijaciona Cox-ova analiza	80
6. DISKUSIJA.....	82
6.1. Efekti segmentacije histograma na prognoštički učinak teksturalnih parametara histopatoloških uzoraka tumora dojke	83
6.1.1. Ekspresija citokeratina u epitelnim malignim ćelijama	83
6.1.2. Učinak u podintervalima niskih i srednjih intenziteta imunobojenja	84
6.1.3. Učinak u podintervalima specifičnog i nespecifičnog imunobojenja	87
6.2. Efekti optimizacije praga binarizacije slika sive skale na prognoštički učinak morfometrijskih parametara epitelnih malignih plaža	89
7. ZAKLJUČAK.....	95
8. LITERATURA	98

1. UVOD

Prema GLOBOCAN bazi podataka čiji podaci obuhvataju istraživanja iz 185 zemalja i uključuju 36 tipova kancera, na globalnom nivou u toku 2020. godine je zabeleženo 19,3 miliona novih i 9,96 miliona smrtnih slučajeva (1). Procene na planetarnom nivou ukazuju da 1 od 5 muškaraca i 1 od 6 žena obole od raka u toku svog života pri čemu je mortalitet u muškoj i u ženskoj populaciji procenjen na 1 od 8 i 1 od 11, respektivno (2).

Rak dojke je jedan od tri najzastupljenija tipa raka među novim slučajevima dok se prema smrtnosti nalazi na petom mestu. U toku 2020 godine je registrovano 2261419 novih slučajeva raka dojke (11,7% ukupne incidencije maligniteta) i 684996 smrtnih slučajeva (6,9% ukupnog mortaliteta usled maligne bolesti). Među ženskom populacijom na globalnom nivou rak dojke je najčešće dijagnostikivan sa 24,5% slučajeva. Takođe je i najčešći uzrok smrti među ženskim pacijentima koji boluju od kancera sa udelom od 15,5% (1).

Rak dojke je bolest sa stalnim porastom učestalosti od oko 3% u poslednje tri decenije. Pogađa sve segmente društva kako u razvijenim tako i u zemljama u razvoju pri čemu stopa novih slučajeva i smrtnih ishoda iskazuje geografsku nejednačenost. Stopa incidencije i mortaliteta je raspodeljena na sledeći način: Azija - 45,4%/50,5%, Evropa - 23,5%/20,7%, Severna Amerika - 12,4%/7,1%, Latinska Amerika i Karibi - 9,3%/8,5%, Afrika - 8,2%/12,5% i Okeanija - 1,1%/0,7% (1). U Republici Srbiji u toku 2020. godine je zabeleženo 6724 novih slučajeva raka dojke što je predstavljalo 28,6% novih slučajeva maligniteta u ženskoj populaciji pri čemu su zabeležena i 2343 smrtna ishoda (3).

Rano dijagnostikovanje i primena terapije kao i stalni napredak u adjuvantnoj terapiji rezultirali su opadanjem stope smrtnosti pacijenata obolelih od raka dojke. Primarni tumor dojke nije smrtonosan nego do smrtnog ishoda dolazi skoro isključivo usled nastanka metastaza čiji se razvoj pokušava sprečiti primenom citotoksične terapije. Međutim, značajan broj pacijenata ne razvije metastaze (4) pa neselektivna primena ovog terapijskog modaliteta dovodi do nepotrebnog izlaganja pacijenata toksičnim efektima (5).

Poboljšanje preživljavanja i kvaliteta života pacijenata može se postići tačnijom prognozom pojave metastaza. U tom cilju se koriste *prognostički faktori* koji omogućavaju klasifikaciju pacijenata na one sa manjim rizikom koji bi mogli da dobiju manje intenzivni tretman i na pacijente visokog rizika za koje se očekuje da će imati koristi od podvrgavanja intenzivnoj terapiji (6). Ovakav medicinski pristup, nazvan *personalizovana medicina*, bi se mogao dostići računarskom, to jest *programski podržanom kvantitativnom analizom* biomedicinskih slika. Između ostalog, to je već ustaljena praksa u slučaju magnetno rezonantnog snimanja, ultrazvučne dijagnostike i pozitronsko emisijne tomografije (7).

Izuzetan napredak računarske tehnologije u kombinaciji sa razvojem tehnika medicinskog snimanja omogućuje brzu i ekonomičnu klasifikaciju i analizu kako medicinskih slika tako i digitalnih slika histopatoloških uzoraka. Efikasnija klasifikacija slika je danas moguća i usled napretka u savremenoj *morfometrijskoj analizi baziranoj na kvantifikaciji i ispitivanju morfoloških karakteristika kao što su broj, veličina i oblik strukturnih elemenata* uz eliminaciju pristrasnosti u proceni rezultata. Računarska analiza biomedicinskih slika u poređenju sa drugim metodama je preciznija, objektivnija, reproduktibilnija i ekonomičnija (8).

Trenutno se u prognozi ishoda raka dojke najčešće koriste hormonski markeri kao i veličine koje su u vezi sa biološkim karakteristikama tumora. Histološki gradus je tradicionalni prognostički faktor koji uzima u obzir histomorfološke karakteristike tumora dojke kao što su zastupljenost tubularnih struktura, pleomorfizma ćelijskih jezgara i mitotičkih ćelija (9). Faktor koji uzima u obzir kliničko-patološke karakteristike tumora se označava kao TNM¹ status, a molekularni prognostički markeri uključuju status steroidnih receptora, amplifikaciju receptora humanog epidermalnog faktora rasta HER2 (10), kao i genske profile kao što su MammaPrint i Oncotype DKS (11,12). Molekularna klasifikacija raka dojke u luminalni A, luminalni B, bazaloidni (trostruko negativni) i HER2+ tip, takođe ima prognostičku vrednost (13).

Prognoza metastaza još uvek nije dovoljno pouzdana jer čak i najnapredniji alati za analizu genskog profila omogućavaju tačnost od samo 65% i površinu ispod

¹Akronim izraza na engleskom *Tumor Node Metastasis*, a odnosi se na veličinu tumora, broj i lokaciju zahvaćenih limfnih čvorova i prisustvo metastaza.

ROC² krive sa postignutom AUC³ vrednošću od 0,69 (14). TNM status je opšte prihvaćen za prognozu ishoda bolesti i određivanje smernica za lečenje kancera (12). Međutim, kod pacijenata bez zahvaćenih limfnih čvorova i bez metastaza nije prognostički upotrebljiv jer kod njih uzima u obzir samo veličinu tumora. Ovo je pokazatelj potrebe za novim prognostičkim faktorima koji bi trebalo da kompenzuju ova ograničenja (15).

Neophodnost poboljšanja prognoze ishoda raka dojke je jak motiv za istraživanja u pravcu kvantifikacije mikroskopske morfologije tumora. Opravdano je pretpostaviti da aktuelni prognostički faktori ne pružaju dovoljnu pouzdanost jer nisu u mogućnosti da u dovoljnoj meri definišu heterogenost tumora. Postojanje intertumorske i intratumorske heterogenosti raka dojke je dokazano u genomskim, histološkim i radiološkim analizama. Računarska analiza histopatoloških uzoraka pokazuje potencijal da pruži precizniji opis tumorske heterogenosti i time poveća prognostičku tačnost. Ovaj vid analize je osmišljen kao strategija koja omogućava da se dospe do informacija koje se ili jako teško ili se nikako ne mogu kvantifikovati vizuelnim mikroskopskim histopatološkim pregledom. Uglavnom se radi o informacijama o prostornoj raspodeli bojenja, distribuciji, obliku i kompleksnosti strukturnih elemenata (16–20). Među prednostima računarske analize su takođe i brzina, efikasnost i ekonomičnost. *Važno je napomenuti da računarska analiza tumorske morfologije ne zaobilazi molekularnu analizu već je integriše tako što kvantifikuje obrasce rasta malignih ćelija koji su upravo posledica ukupnih genetskih modifikacija i molekularnih interakcija unutar tumora.*

Ovakav pristup dodatno dobija na značajnosti usled napretka računarske brzine kao i dostupnosti digitalnog skeniranja histopatoloških uzoraka (21).

Prethodne studije su ukazale na prognostički značaj računarske analize tumorske morfologije kod raka dojke (4,22–24). Međutim, korišćene matematičke veličine su kompleksne, vrlo apstraktne i komplikovane za interpretaciju. Zato je cilj

²Akronim izraza na engleskom *Receiver Operating Characteristic* ili na srpskom, *radna karakteristika prijemnika*; označava metod grafičke analize efikasnosti diskriminatora zasnovan na nalaženju kompromisa između senzitivnosti i specifičnosti.

³Akronim izraza na engleskom *Area Under Curve* ili na srpskom, *površina ispod krive*, a u medicinskim naukama se odnosi na površinu ispod ROC krive.

ove teze da se ispita prognostička efikasnost jednostavnih morfoloških karakteristika kao što su veličina, oblik i intenzitet bojenja klastera malignih epitelnih ćelija, takozvanih *epitelnih malignih plaža* (engl. *malignant nests*) (25), u histopatološkim uzorcima raka dojke. Epitelne ćelije formiraju epitelno tkivo koje čini zaštitnu granicu svih spoljašnjih i unutrašnjih telesnih površina i ulazi u sastav žlezda. Epitelno tkivo je najčešće podložno malignoj transformaciji pri čemu kancerima epitelnog porekla, takozvanim karcinomima, pripada skoro 90% maligniteta kod ljudi (26).

1.1. Prognostički faktori ishoda primarnog kancera dojke

Prognostički faktori su povezani sa prirodnim tokom i ishodom bolesti bez obzira na terapiju te su stoga značajni za predviđanje razvoja bolesti (27). Prognostički faktori su diskriminatori koji se koriste za klasifikaciju pacijenata u funkciji ishoda bolesti. Po svojoj prirodi mogu biti kontinuirane ili diskretne veličine. Treba ih razlikovati od prediktivnih faktora koji koreliraju sa odgovorom pacijenta na specifičnu terapiju jer odražavaju senzitivnost tumora na određeni vid tretmana. Kako u cilju prognoze tako i u cilju predikcije koriste se faktori koji su po svojoj prirodi međusobno različiti. Uglavnom se koristi kombinacija karakteristika pacijenata i objektivno merljivih veličina koje su u vezi sa biološkim osobinama tumora.

Neki faktori imaju i prognostičku i prediktivnu vrednost. Klasični primeri su hormonski receptori i povećana ekspresija HER2 (28).

Kvalitetan prognostički faktor može da ustanovi razliku u verovatnoći ishoda bolesti. Zavisno od svoje pouzdanosti, prognostički faktor se može kvalifikovati kao slab, umeren ili jak, ili kao neprihvatljiv, prihvatljiv, dobar, itd. (29). Status limfnih čvorova je primer faktora sa visokim stepenom pouzdanosti. Kod pacijenata kod kojih je ustanovljeno prisustvo malignih ćelija u limfnim čvorovima je verovatnoća recidiva je dva do tri puta veća. Primer prognostičkog faktora sa nešto slabijom pouzdanošću je ekspresija estrogenskih receptora. ER+ pacijenti imaju lošije ishode od ER- pacijenata i zato je ekspresija ER jak prediktivni faktor za hormonsku terapiju sa mogućim smanjenjem rizika od recidiva i do 50%. (30).

Statistički značaj nekog prognostičkog faktora ne garantuje i njegov klinički značaj. Mnogi nezavisni faktori mogu dobro da koreliraju sa ishodom bolesti ali podvrgnuti multifaktorskoj analizi mogu da izgube statističku značajnost (29). Čak i kada zadrži statistički značaj u multifaktorskoj analizi prognostički faktor može da bude od neznatne kliničke vrednosti. Moguće je da biološki značajan faktor bude od minimalne vrednosti u proceni verovatnoće nekog događaja. Ta je situacija tipična kada je odabrani faktor relativno retko prisutan u grupi ispitivanih pacijenata što ograničava njegovu upotrebnu vrednost.

Jedan od zahteva za odabir i upotrebu prognostičkog faktora je njegova reproduktibilnost i tehnička pouzdanost. Pouzdanost je obrnuto proporcionalna varijabilnosti u njegovoj interpretaciji. Instrukivan primer je već spomenuta gradacija tumora koja je i pored značajnog poboljšanja algoritma gradacije neizostavno podložna varijacijama u interpretaciji od strane patologa. Pored toga, razlike u ocenjivanju između laboratorija ili razlike u detaljima testa dovode do varijabilnosti u tumačenju rezultata čak i standardnih biohemijskih tumor markera, na primer u tumačenju ekspresije HER2 (31).

Optimalna prognoza se postiže analizom svih dostupnih prognostičkih faktora. U današnje vreme je značajna primena prognostičkih indeksa koji omogućavaju procenu rizika uključivanjem nekoliko kliničkih parametara. Među najpopularnijim i najčešće korišćenim indeksima su Adjuvant!, PREDICT i Oncotype DX® (32).

1.1.1. Heterogenost tumora dojke i standardno korišćeni prognostički faktori

Jedno od bitnih obeležja raka dojke je njegova inter i intratumorska heterogenost. Velike razlike u karakteristikama tumora između pacijenata ali i unutar pojedinačnih tumora predstavljaju veliki dijagnostički, prediktivni i prognostički izazov. Razumevanje mehanizama i faktora koji dovode do heterogenosti tumora je od izuzetne važnosti za razumevanje razvoja bolesti i za odabir najefikasnijeg terapijskog modaliteta. Uglavnom su prepoznata četiri mehanizma heterogenosti tumora (33): nivo diferenciranosti ćelije porekla, međusobna konverzija između CSC (engl. cancer stem cells) i ne-CSC koja određuje

plastičnost ćelija tumora, genetska evolucija tumora i interakcija sa mikro okruženjem koja može da utiče pozitivno ili negativno na karcinogenezu.

Heterogenost tumora pruža informacije od izuzetnog značaja za prognozu ishoda bolesti. Među standardno korišćenim prognostičkim faktorima ishoda kancera dojke izdvajaju se dve klase: karakteristike povezane sa biološkim osobinama tumora i opšte karakteristike pacijenata. Modeli razvijani u poslednjim decenijama su uglavnom bazirani na korišćenju limfno-nodalnog statusa, veličine i gradusa tumora, starosti pacijenta u trenutku dijagnoze i ER statusa (34).

Najčešće dostupni i korišćeni faktori, od kojih su većina razmatrani za potrebe ove teze, mogu se identifikovati kao:

1- *Karakteristike domaćina tumora.* Ove karakteristike uključuju: starost, status menopauze i etničku pripadnost. Studije koje su razmatrale prognostički učinak starosti na ishod bolesti su uglavnom ukazale na nedovoljnu pouzdanost ovih faktora. Ipak, analize koje su uključivale i multifaktorsku analizu su ukazale da starost ispod 35 godina ukazuje na nešto lošiji ishod bolesti (35,36).

Iako se etnička pripadnost i starost mogu koristiti kao prognostički faktori, preporučuje se njihova kombinovana upotreba sa pouzdanijim faktorima, kao što je, na primer, veličina tumora.

2- *Biološke karakteristike tumora* se koriste za analizu i kvantifikaciju kliničke i histopatološke heterogenosti tumora i uglavnom su primarni razlog različitog ishoda bolesti pacijenata koji boluju od raka dojke (33). Najčešće korišćeni faktori za opis i kvantifikaciju morfološke heterogenosti su TNM status i gradus tumora.

- *Aksilarni limfo-nodalni status.* Za pacijente sa rakom dojke dijagnostikovanim u ranoj fazi najznačajniji prognostički faktor je aksilarni nodalni status koji kvantifikuje broj zahvaćenih limfnih čvorova. Još davno je ukazano na postojanje direktne veze između rizika za pojavu udaljenih metastaza i broja zahvaćenih aksilarnih limfnih čvorova (37,38). Uobičajeno je da se za kliničke svrhe pacijenti prema limfno-nodalnom statusu svrstaju u jednu od četiri grupe: prva (bez zahvaćenih čvorova), druga, treća i četvrta sa 1 - 3, 4 - 9 i 10 ili više zahvaćenih limfnih čvorova, respektivno (39). Petogodišnja stopa preživljavanja u zavisnosti od broja zahvaćenih čvorova je: 28,4% (13 ili više

čvorova), 45,7% (4 - 12 čvorova), 73% (1 - 3 čvorova) i 82,8% za pacijentkinje sa negativnim nalazom (40). Podaci ukazuju da je pozitivan nalaz preporuka za adjuvantnu sistemsku terapiju zato što se smatra da je rizik za povratak bolesti veći od 20% dovoljan da se opravdaju neželjeni efekti ovakvog vida terapije. Za pacijente niskog rizika sa negativnim nalazom je neophodna individualna prognoza bazirana na drugim prognostičkim faktorima.

- *Veličina tumora.* Najveća dimenzija tumora je u korelaciji sa brojem zahvaćenih limfnih čvorova ali se koristi i kao nezavistan prognostički faktor. Broj udaljenih metastaza i verovatnoća smrtnog ishoda su direktno proporcionalni veličini tumora (38,41). Veličina tumora kod pacijenta sa pozitivnim nalazom za limfno-nodalni status nije nezavistan prognostički faktor. Međutim, kod pacijenata sa negativnim nodalnim nalazom ovo je najbolji prognostički faktor koji se koristi za odluku o adjuvantnoj terapiji. Pacijenti sa tumorom veći od 1-2 cm su preporučeni za adjuvantnu terapiju zato što je rizik od recidiva veći od 20%.
- *Tip i histološki gradus tumora.* Oba ova faktora se rutinski koriste za izbor terapije i za prognozu razvoja bolesti. Najčešće korišćeni metod određivanja histološkog gradusa je Nottingham sistem gradiranja (42), tj. Elston-Ellis modifikacija SBR (Scarff-Bloom-Richardson) sistema (43) po kome se gradus izračunava na osnovu tri faktora: zastupljenost tubularnih struktura, stepen pleomorfizma i broj mitoz. Ovim faktorima se dodeljuju vrednosti od 1 do 3 koje se sabiraju i koriste u određivanju gradusa: G1 tako ima zbir od 3 - 5 i odgovara dobro diferenciranim tumorima, G2 ima zbir 6 ili 7 i odgovara umereno diferenciranim tumorima dok G3 ima zbir 8 ili 9 i odgovara slabo diferenciranim tumorima. Za histološki gradus je pokazano da ne zavisi od limfno-nodalnog statusa i veličine tumora (42) i da je relativno pouzdan prognostički faktor za ukupno preživljavanje pacijenata (39). Pri tome je nađeno i da je rizik od recidiva veći kod pacijenata sa visoko gradiranim tumorima (44). Kombinovana upotreba gradusa sa nodalnim statusom i veličinom tumora se pokazala kao pouzdaniji metod od pojedinačnih faktora (34).

3- *Biohemijski hormonski faktori* su povezani sa biološkim karakteristikama tumora ali se obično identifikuju kao posebna grupa faktora jer se koriste i u prognozi i u predikciji. Kao biomarkeri tumorske heterogenosti se uglavnom koriste status estrogenskih i progesteronskih receptora (ER i PR, respektivno) i HER2 status pri čemu je njihova ekspresija jedan od kritičnih elemenata pri odlučivanju o izboru terapije (45). Velike varijacije ekspresije biomarkera su uočene u okviru istog tumora (46), pri čemu nivo ekspresije korelira se efikasnošću terapije (30,47).

- *ER/PR status.* Približno 80% pacijenata koji boluju od karcinoma dojke su ER+, dok je procenat PR+ pacijenata između 60 i 70% (30,47). Status estrogenskih i progesteronskih receptora u analizi raka dojke se koristi kako za predikciju ishoda terapije tako i za prognozu razvoja bolesti. Optimalna primena ER/PR statusa je u predikciji kliničke korisnosti adjuvantne terapije. Između 70 i 80% pacijentkinja koje boluju od raka dojke su ER+/PR+ (33). Ove pacijentkinje se tretiraju hormonskom terapijom kao što je tamoksifen, pri čemu se kod približno 60% ovih pacijentkinja dobija pozitivan odgovor na terapiju (9). ER+/PR- i ER-/PR+ pacijenti su vrlo retki.

Prognostička primena ER i/ili PR statusa je ograničena. Njihov prognostički učinak je od značaja kod hirurški tretiranih pacijenata (6). Na prognostički značaj ER/PR statusa ukazuju podaci iz NSABP (engl. National Surgical Adjuvant Breast Project) studije koja je obuhvatila pacijentkinje sa rano otkrivenim i hirurški odstranjenim tumorom dojke, bez primene adjuvantne terapije (48). ER+ pacijenti su u prvih pet godina praćenja imali stopu preživljavanja bez bolesti (DFS⁴) od 74% i ukupnu stopu preživljavanja (OS⁵) od 92%, dok je za ER- pacijente nađeno DFS = 66% i OS = 82% (6,29). ER status je dakle slab prognostički faktor pri čemu i kada se kombinuje se PR statusom ne dolazi do poboljšanja učinka.

- *HER2/neu status.* Amplifikacija HER2 receptora se detektuje u oko 20% slučajeva kancera dojke (28). Inače je ekspresija niskog nivoa ovog faktora prisutna i kod normalnih epitelnih ćelija. Visok HER2 status je indikator lošije prognoze naročito kod pacijenata sa pozitivnim limfno-nodalnim nalazom

⁴ Akronim izraza na engleskom "Disease-Free Survival", ili na srpskom, *stopa preživljavanja bez bolesti*, definisana kao vremenski interval između tretmana i ponovne pojave maligniteta.

⁵ Akronim izraza na engleskom "Overall Survival", ili na srpskom, *ukupna stopa preživljavanja*.

(28,49). Istovremeno je HER2 amplifikacija jak prediktivni faktor za upotrebu HER2 inhibitora herceptina.

1.2. Računarska analiza digitalnih slika histopatoloških uzoraka raka dojke

Tradicionalni prilaz morfološkoj analizi je ustanovljen još početkom prošlog veka i uglavnom se oslanja na vizuelnu analizu i subjektivni kvalitativni opis. Poslednje decenije 20. veka su obeležene izuzetnim napretkom u kvantitativnoj analizi ćelijskih klastera, pojedinačnih ćelija i substrukturnih jedinica. Ovaj napredak je povezan sa napretkom računarske tehnologije, morfometrijskih deskriptora i programski podržanih alata koji omogućavaju morfološke studije bazirane na računarskoj analizi slika. Potreba za objektivizacijom morfološke analize proizilazi i iz nedostatka standardizovanih metoda i parametara, što je bitno za izvođenje komparativnih studija. Savremena morfometrijska ćelijska analiza omogućava računanje parametara koji se koriste kao deskriptori specifičnih geometrijskih i teksturalnih karakteristika (50).

Morfometrijski parametri su objektivno merljive morfološke karakteristike koje mogu biti definisane kao oblik, veličina, intenzitet i tekstura, a koje su na direktan ili indirektan način povezani sa njihovim biološkim svojstvima. Morfometrijski parametri su primenljivi u analizi bioloških struktura i ukazuju na osobine koje se teško uočavaju i/ili kvantifikuju mikroskopski pregledom (50). Računarska analiza digitalnih slika histopatoloških uzoraka je dopuna vizuelnoj analizi patologa jer kvantifikuje ono što patolog ne može vizuelno da definiše.

1.2.1. Digitalna patologija i računarski prognostički faktori

Standardno korišćeni prognostički faktori ne obezbeđuju dovoljnu pouzdanost u prognozi pojave metastaza (15) verovatno usled toga što ne iskorišćavaju u potpunosti heterogenost tumora i morfometrijske karakteristike pripadajućih strukturnih elemenata. Kao što je već napomenuto, histološki gradus uzima u obzir obrasce rasta malignih ćelija koji su posledica molekularnih interakcija unutar tumora i zato je bogat

izvor informacija povezanih sa njegovom heterogenošću, ali je ipak nedovoljnog prognostičkog kapaciteta. Dostupno rešenje za prevazilaženje ograničenja standardnog pristupa i poboljšanja kvaliteta prognoze je u *digitalnoj patologiji* koja se bavi prikupljanjem, skladištenjem i interpretacijom morfoloških informacija koristeći digitalizovane slike patoloških uzoraka (51).

Nezaobilazni element digitalne patologije je prikupljanje slika u digitalnom formatu koje se razvijalo od upotrebe digitalnih kamera instaliranih na mikroskopu do modernih skenera koji omogućavaju digitalizaciju celokupnih histoloških preparata (52). Digitalni format olakšava kako kvalitativnu tako i kvantitativnu analizu uzorka (53) oslanjajući se na algoritme matematičke i statističke analize. Dok je tradicionalni pristup analizi histopatoloških uzoraka uglavnom deskriptivne prirode i baziran na subjektivnoj proceni, programski podržana analiza računa veliki broj objektivno određenih kvantitativnih parametara koji su odraz morfoloških karakteristika uzorka (54,55).

Više istraživanja je ukazalo na to da se rezultati računarske analize histopatoloških uzoraka po svojoj pouzdanosti i preciznosti mogu porediti sa rezultatima mikroskopske analize (56–58). Vizuelna analiza patologa je ipak i dalje nezamenljiva za odabir reprezentativnih preseka tumora koji se kasnije podvrgavaju računarskoj analizi.

Za ljudsko oko i mozak, prepoznavanje, interpretacija i razumevanje vizuelnih prikaza, čak i onih koji sadrže izuzetno veliki broj informacija, je uglavnom rutinski zadatak. U slučaju računara, prepoznavanje, interpretacija, klasifikovanje i razumevanje kompleksnih slika, pojedinačnih objekata i elemenata koje sadrže, može da bude izuzetno komplikovano i pored njihovog sve većeg kapaciteta i brzine. Metode analize digitalnih slika korišćenjem niza kvantitativnih karakteristika taj proces olakšavaju, ubrzavaju, automatizuju, a vrlo često omogućavaju i identifikaciju osobina koje izmiču direktnom vizuelnom pregledu.

Računarska analiza medicinskih slika dobijenih tomografskim metodama već je ustaljena praksa (7), dok je u histopatologiji još uvek predominantno direktno vizuelno ispitivanje za koje je neophodno dobro trenirano i iskusno medicinsko osoblje. Takva situacija je uglavnom rezultat tri faktora: (i) izražena varijabilnost histopatoloških

slika usled mešanja morfoloških osobina na jednoj istoj slici, (ii) odsustvo standardizovanih tehnika fotografisanja (55), kao i (iii) odsustvo standardizovanih metoda morfometrijske analize. Aktuelni kapacitet računara, razvoj mikroskopskih tehnika i unapređenja u skeniranju celih histopatoloških uzoraka, omogućavaju brzu i pouzdanu kvantifikaciju i analizu velikog broja morfoloških karakteristika uzoraka.

Primarni tumor je polazište metastatske diseminacije i dragocen je izvor prognostički značajnih informacija (59), pri čemu su udaljene metastaze glavni uzrok smrtnog ishoda bolesti i kao takve, najvažniji događaj u prognozi individualnog ishoda raka dojke. Prevalencija karcinoma, kancera čije je poreklo u epitelnim ćelijama dojke, je verovatno posledica visoke učestalosti deobe epitelnih ćelija. Epitelne ćelije, kako normalne tako i maligne se mogu efikasno bojiti monoklonskim antitelima AE1 i AE2 koja su dobro definisana i poseduju specifičnost za keratinska vlakna (60). Maligne epitelne ćelije su uglavnom grupisane u takozvanim tumorskim plažama različitih veličina i oblika, rasutim unutar strome zajedno sa fibroblastima, imunim ćelijama i nekolicinom epitelnih izdanaka (15). U tom kontekstu je analiza epitelnih malignih plaža vrlo dragocena jer su epitelne ćelije izvor neoplastičnih transformacija u gotovo 99% slučajeva raka dojke (61).

Uobičajeno je da takvi strukturni elementi iskazuju veliku varijabilnost kako u veličini tako i u obliku između različitih tumora, ali i unutar jednog istog tumora. Ta činjenica je dovela do pretpostavke da su oblik i veličina malignih klastera identifikovanih u slikama histopatoloških uzoraka potencijalno bogat izvor prognostički korisnih informacija i da mogu da se koriste u proceni metastatske agresivnosti tumora. Zajedno sa prethodnim, rezultati koji su ukazali da je niski intenzitet bojenja pan-citokeratinom povezan sa lošijim ishodom (62) su motiv za detaljnu računarsku analizu prognostičke vrednosti informacija povezanih sa prostornom raspodelom intenziteta pan-citokeratinskog bojenja (61) na digitalnim slikama histopatoloških uzoraka raka dojke.

Algoritmi uobičajeno korišćeni za računarsku analizu medicinskih slika se mogu grupisati na sledeći način: statistički (npr., statistike prvog i drugog reda), strukturni (npr. Voroni-jevi poligoni, ivice), lokalni binarni obrasci (npr. LBP, CLBP), bazirani na modelu (npr. autoregresivni model, model Markov-ljevih

nasumičnih polja, fraktali) i transformacije (npr. Gaborova i wavelet transformacija, prostorni i frekventni filteri) (63). Ovi algoritmi su korišćeni za analizu nespecifično (23) i specifično (62,64) bojenih histopatoloških uzoraka. Oni su vrlo dobrog prognostičkog kvaliteta sa AUC koji postiže vrednosti do 0,77 (22). Međutim, njihovo računanje je komplikovano i time su teksturalne karakteristike koje se detektuju vrlo apstraktne i zato se teško dovode u direktnu vezu sa vizuelno prepoznatljivom morfologijom histopatoloških uzoraka. Pomenute poteškoće su najverovatniji razlog inercije u prihvatanju računarskih parametara od strane medicinskih radnika i njihovoj integraciji u standardnu kliničku praksu. Deo analiza vršenih za potrebe ove disertacije bi mogao da olakša prevazilaženje tih problema jer se bazira na analizi morfoloških parametara epitelnih malignih plaža koji su jednostavni za interpretaciju (15).

2. HIPOTEZE ISTRAŽIVANJA

Optimizacija računarske analize morfologije epitelnih malignih plaža tumora dojke u ovoj doktorskoj tezi je zasnovana na proveru sledećih hipoteza:

1. Podela histograma intenziteta piksela digitalnih slika histopatoloških uzoraka raka dojke može da unapredi prognostički učinak teksturalnih parametara.
2. Optimizacija praga binarizacije digitalnih slika histopatoloških uzoraka raka dojke može da unapredi prognostički učinak morfometrijskih parametara.
3. Prognostički učinak morfometrijskih i teksturalnih parametara digitalnih slika histopatoloških uzoraka raka dojke pokazuje uspešnije rezultate u odnosu na prognostički učinak kliničko-patoloških parametara tumora.

3. CILJEVI ISTRAŽIVANJA

Provera postavljenih hipoteza je izvršena ostvarenjem ciljeva istraživanja koji se odnose na optimizaciju pripreme digitalnih slika histopatoloških uzoraka raka dojke, izračunavanje morfometrijskih i teksturalnih parametara, procenu njihovog prognostičkog učinka i poređenje sa učinkom standardno korišćenih prognostičkih faktora.

Prva hipoteza se proverava ostvarenjem sledećih ciljeva:

- ispituje se raspodela intenziteta pan-citokeratinskog bojenja na osnovu čega se histogram intenziteta deli na nepreklapajuće podintervale,
- na slikama segmentiranim prema ustanovljenim podintervalima se ispituje prognostički učinak teksturalnih parametara (GLCM i parametara statistike prvog reda) primenom ROC analize,
- analizira se efekat podele intervala intenziteta pan-citokeratinskog bojenja na prognostički učinak teksturalnih parametara.

Druga hipoteza se proverava ostvarenjem sledećih ciljeva:

- procesom automatske binarizacije i primenom korisnički odabranih pragova binarizacije, izdvajaju se slike malignih plaža različitog intenziteta pan-citokeratinskog bojenja i analizira se prognostički učinak morfometrijskih parametara (broja, veličine i oblika plaža) primenom ROC analize,
- analizira se efekat primene različitih pragova binarizacije na prognostički učinak morfometrijskih parametara,
- ispituje se raspodela veličine i cirkularnosti epitelnih malignih plaža na osnovu čega se interval vrednosti veličine i cirkularnosti deli na nepreklapajuće podintervale,
- na binarizovanim slikama segmentiranim prema ustanovljenim podintervalima veličine i cirkularnosti se ispituje prognostički učinak morfometrijskih parametara primenom ROC analize,
- analizira se efekat primene različitih filtera veličine i cirkularnosti na prognostički učinak morfometrijskih parametara.

Treća hipoteza se proverava ostvarivanjem sledećih ciljeva:

- ispituje se prognostički učinak demografskih i kliničko-patoloških prognostičkih parametara analizirane grupe pacijentkinja primenom ROC analize,
- poredi se prognostički učinak demografskih i kliničko-patoloških sa učinkom teksturalnih i morfometrijskih parametara.

4. MATERIJALI I METODE

4.1. Uzorak pacijenata

Grupu analiziranih pacijenata čine 102 pacijenta ženskog pola, kavkaskog tipa, starosti koje se u trenutku dijagnostikovanja 1993. godine kretalo između 37 i 80 godina (srednje vrednosti od 57 godina). Pacijentkinje su na osnovu demografskih i kliničko-patološki karakteristika klasifikovane kao pacijenti niskog rizika i tretirane su u skladu sa preporukama važećim u trenutku uspostavljanja dijagnoze. Sve pacijentkinje su podvrgnute adekvatnom hirurškom zahvatu i tretirane su zračenjem.

Prospektivni proračun veličine uzorka (MedCalc Software, Ostend, Belgium) je izvršen na osnovu pilot studije koja je uključivala 40 pacijentkinja sa vrednostima $\alpha = 0,05$, $\beta = 0,20$ i sa veličinom efekta izraženom koristeći AUC vrednost od 0,72, tj. 0,28. Proračun je ukazao da bi uzorak trebao da sadrži 85 pacijenata od kojih su 17 pozitivni. Korišćeni uzorak se sastoji od 102 pacijentkinje od kojih su 21 bili pozitivne, pri čemu je maksimalna dobijena AUC vrednost iznosila 0,82. Uzorak je odabran retrospektivno odabirom samo onih pacijentkinja koje nisu podvrgnute sistemskom lečenju baziranom na primeni hormonskih i citotoksičnih preparata.

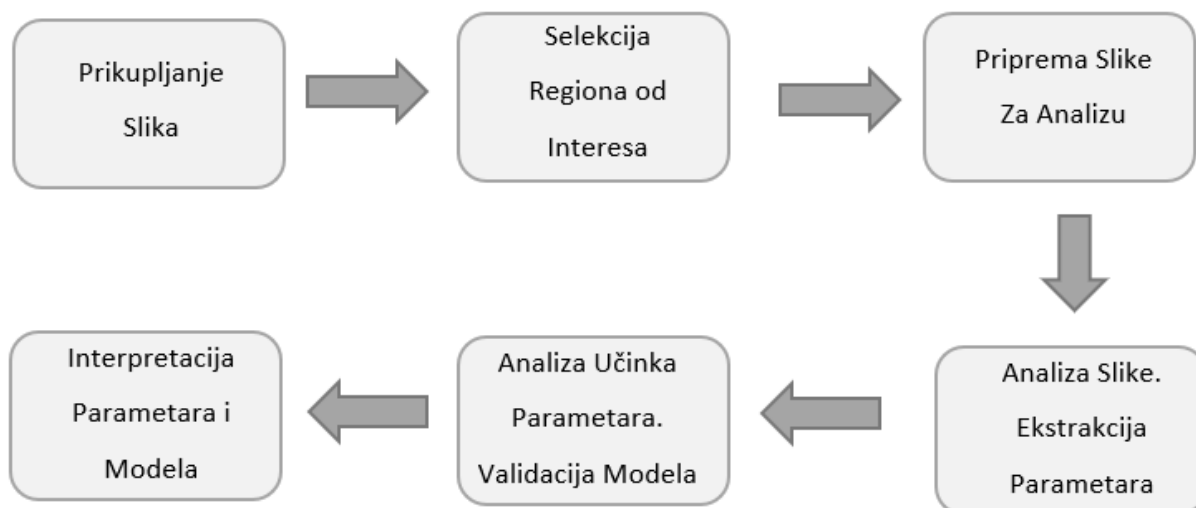
Ova studija (15,61) je odobrena od strane Etičkog odbora Instituta za Onkologiju i Radiologiju Srbije, broj dokumenta: 01-972, 07. april 2021 godine.

4.1.1. Testiranje HER2 amplifikacije

U periodu uspostavljanja dijagnoze, 1993. godine, procena HER2 amplifikacije nije bila uključena u rutinska klinička ispitivanja. Ove analize su obavljena naknadno, 2012 godine, korišćenjem hromogenske *in situ* hibridizacije CISH. Upotrebljen je SPOT-Light® HER2 CISH kit (Kat. br. 84-0150, Zymed/Invitrogen, Thermo Fisher Scientific Corp. Waltham, MA, USA) kao što je opisano detaljno od strane Todorović-Raković et al. (65). Od 2000 godine CISH tehnike se upotrebljavaju kao alternativa FISH tehnikama utemeljenim na fluorescentnoj *in situ* hibridizaciji

(66). CISH se bazira da kolorimetrijskoj detekciji samo jedne probe dok FISH koristi dve boje. U odnosu na fluorescentne metode glavna mu je prednost što se ne bazira na fluorescenciji tako da se rezultati očitavaju na standardnom optičkom mikroskopu (67). U izvršenim analizama je korišćena DNK proba obeležena digoksinom. Za procenu rezultata hibridizacije je korišćen Olympus BX51 mikroskop pri uvećanjima od 400x i 1000x. Prihvaćena granična vrednost koja definiše HER2 amplifikaciju kao pozitivnu je šest ili više kopija gena po jezgru. Od ranije je poznata relativno visoka koincidencija rezultata procene amplifikacije HER2 gena imunohistohemijskim i CISH metodom (65,66).

4.2. Protokol rada u računarskoj analizi digitalnih slika



Slika 4.2.1. Uobičajeni protokol rada u kvantitativnoj analizi digitalnih slika histopatoloških uzoraka

Klasični protokol rada u računarskoj analizi digitalnih biomedicinskih slika je višefazni proces (slika 4.2.1). Prvi korak je prikupljanje slika. U slučaju histopatoloških uzoraka se koriste fotografije uzoraka ili slike dobijene skeniranjem celih histoloških preparata upotrebom skenera. U nastavku se od strane stručnog lica vrši odabir najreprezentativnijih oblasti slike za koje se smatra da sadrže najznačajnije histopatološke informacije. U sledećoj fazi se vrši priprema slika za dalju analizu, kao na primer, razdvajanje kanala boja, normalizacija i binarizacija. Preprocesovane slike se zatim analiziraju (matematički i statistički) čime se generiše veliki broj parametara i rezultata. Dobijeni parametri i rezultati se zatim statistički

porede sa ishodom bolesti čime se procenjuje njihova klinička upotrebna vrednost, nezavisnost i korelacija, između ostalog. Nakon izvršene procene učinka i značajnosti, odabira se reprezentativan skup parametara koji se na kraju podvrgavaju tumačenju u biološko-medicinskom i u kontekstu kliničke primenljivosti.

4.3. Prikupljanje slika i selekcija regiona od interesa

Za potrebe ove disertacije analizirane su 532 digitalne slike histopatoloških uzoraka raka dojke. Uzorci su bojeni pan-citokeratinom čime je omogućeno obeležavanje i vizuelizacija klastera malignih epitelnih ćelija.

4.3.1. Histopatološka priprema i bojenje uzoraka

Uzorci tkiva za analizu su prikupljeni u toku hirurškog odstranjivanja tumorskog tkiva. Primarno tumorsko tkivo je fiksirano u formalinu, zatopljeno je u parafin a zatim je sečeno na debljinu od 4 μm . Ovi preseki su nanoseni na staklene pločice.

U presecima tkiva fiksiranim formalinom antigeni su toplotno renaturisani u vodenom kupatilu na 95° C u EDTA pH8 puferu tokom 40 min. Inaktivacija endogene peroksidaze je vršena tokom 30 min korišćenjem 3% H₂O₂ u metanolu. Za preinkubaciju je iskorišćen 5% kozji serum. Proces inkubacije kompletnih tkivnih uzoraka je prvobitno obavljen monoklonalnim zečijim antitelima CD8 (ThermoFisher Scientific, Valtham, MA, SAD; #RM-9116-S1) a zatim i klonovima primarnih pan-citokeratinskih antitela mAE1/AE3 (Dako, Glostrup, Danska, #M3515) u 5% kozijem serumu tokom 60 min. Ovako primenjen koktel antitela boji epitelne ćelije tako što detektuje ukupno 13 citokeratina: 1 – 8, 10, 14 – 16 i 19, pri čemu je CD8 izvor braon tonova, dok su pan-citokeratinska antitela izvor plavih tonaliteta (15).

Ispiranje je izvršeno u PBS pri čemu je dodat i sekundarni kozji anti-zečiji IgG HRP konjugat (Jackson ImmunoResearch Laboratories, Vest Grove, PA, USA; #111-035-144). Nakon toga je dodata alkalna fosfataza konjugovana sa poliklonalnim kozjim anti-mišijim IgG (Southern Biotech, Birmingham, AL, USA; #1030-04) u 5%

kožjem serumu. Hromogeni su pojačani prvo DAB niklom (Vector Laboratories, Burlingame, CA, USA) a zatim i sa FastBlue RR (Sigma-Aldrich, St. Louis, MO, USA). Nije korišćeno kontrastno bojenje zato da bi se istakle samo epitelne ćelije (slika 4.4.1 a).

4.3.2. Skeniranje uzoraka i izdvajanje reprezentativnih segmenata

Digitalne dvobojne plavo-braon slike (slika 4.4.1 a) su dobijene skeniranjem histopatoloških uzoraka korišćenjem digitalnog skenera slajdova visoke rezolucije NanoZoomer Hamamatsu-XRC1200 (Hamamatsu Photonics K.K., Japan). Dobijene su slike visoke rezolucije na kojima su prikazani celokupni uzorci tkiva.

Zarad postizanja maksimalne reprezentativnosti i pouzdanosti, umesto slike celokupnog isečka patolog je odabrao segmente koji su sadržavali karakteristične obrasce rasta individualnih tumora, sa najvećom količinom pan-citokeratinski obojenih malignih ćelija i bez prisustva artefakata. Klasteri obojenih epitelnih ćelija u odabranim presecima su procenjeni kao normalni ili kao maligni zavisno od prepoznatih morfoloških karakteristika.

Analize urađene za potrebe ove teze su vršene na plavim slikama dobijenim nakon razdvajanja kanal boja na plavi (pan-citokeratinski) i braon (CD8) izvršeno procedurom koju su detaljno opisali Li i Plataniotis (68).

4.4. Priprema slika za kvantitativnu analizu

Konverzija vrste slike i njene rezolucije zavisila je od primenjene metode analize i kvaliteta rezultata. Bez obzira na metod konverzije određena količina informacije se gubi što ne mora da bude limitirajući faktor ukoliko izgubljena informacija nije u direktnoj vezi sa osobinama koje se žele kvantifikovati.

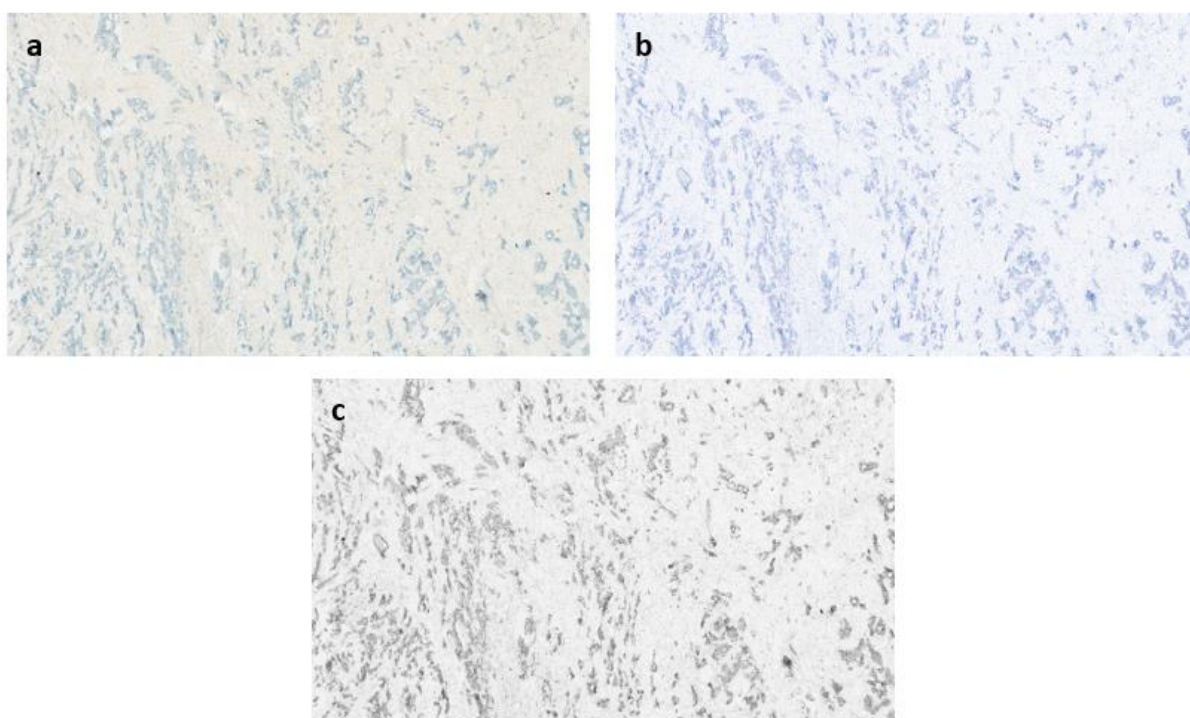
Različite metode analize zahtevaju različit kvalitet i tip slika. Za analize izvedene za potrebe ove disertacije bilo je neophodno izvršiti prilagođavanje formata slike. Računanje teksturalnih parametara je vršena na 8-bitnim slikama sive skale dok je za potrebe morfometrijske analize broja i karakteristika epitelnih malignih plaža, izvršena binarizacija sive skale.

4.4.1. Transformacija originalnih slika u slike sive skale i podela histograma

Konverzija plavih slika u 8-bitne slike sive skale je izvršena ponderisanom konverzijom (69) pri čemu je intenzitet konvertovanih piksela i_{gray} izračunat primenom formule

$$i_{gray} = 0,299red + 0,587green + 0,114blue.$$

Plave slike su transformisane u 8-bitne slike sive skale (slika 4.4.1 b i c) korišćenjem komande "8-bit" (*Image<Type<8-bit*) pri čemu su crnoj i beloj boji dodeljene vrednosti intenziteta 0 i 255 (61), respektivno. Za konverziju je korišćen javno dostupni programa za analizu i obradu slika *ImageJ* (National Institutes of Health, USA, <http://imagej.nih.gov/ij/>) (70).



Slika 4.4.1. (a) Dvobojna slika pan-citokeratinskog imunobojenja histopatološkog uzorka. (b) Slika plave boje dobijena nakon razdvajanja kanala i (c) njena 8-bitna slika sive skale raspona od 0 (crno) do 255 (belo).

Podela celokupnog intervala intenziteta piksela na sedam podintervala je izvršeno korišćenjem programa Fiji/ImageJ verzija 1.53c (71). Izdvajanje željenog

podintervala intenziteta je urađena u dvostepenom procesu pri čemu su svi pikseli intenziteta van odabranog podintervala konvertovani u bele. Na primer, za interval 200–220 u prvom koraku su svi pikseli intenziteta manji od 200 konvertovani u 255, a zatim, u drugom koraku, isto je učinjeno i za sve piksele intenziteta većih od 220. Korišćen je makro: `run (“Macro...”, “code = [if (v >= 0 && v <= 200) v = 255]”)` u prvom, a zatim makro: `run (“Macro...”, “code = [if (v >= 220 && v <= 255) v = 255]”)` u drugom koraku (61). Ovom metodom su dobijeni sledeći podintervali intenziteta piksela: 0 - 130, 130 - 160, 160 - 180, 180 - 200, 200 - 220, 220 - 240 i 240 - 255.

4.4.2. Binarizacija slika

Za potrebe morfometrijske analize su korišćene binarizovane slike (15). Proces binarizacije omogućava segmentaciju slike tako što se pikseli pretvaraju u crne ili u bele poređenjem njihovog intenziteta f_{ij} sa odabranim pragom t . Ako je $f_{ij} \geq t$, piksel (i, j) se konvertuje u beli. U suprotnom se konvertuje u crni. Ovim procesom se slike transformišu u crno-bele tako što se celokupan interval sivih tonaliteta redukuje na bele i crne piksele. Uobičajeno je da beli pikseli odgovaraju pozadini a crni objektu od interesa. Ovakav način segmentacije slike je naročito prigodan kada su od interesa karakteristike analiziranog objekta koje direktno zavise od intenziteta kao što su oblik i prostorna raspodela. To je upravo slučaj sa digitalnim slikama histopatoloških uzoraka kada je primarni zadatak segmentacije slike razdvajanje obojenih ćelijskih klastera od pozadine.

Važnost procesa binarizacije za analize koje se vrše na crno-belim slikama je očigledna pri čemu je odabir praga za binarizacije kritičan korak. Za odabir praga binarizacije razvijeno je više metoda (72,73) od kojih su najdirektnije i najintuitivnije bazirane na histogramu. Matematički kompleksnije metode izbora praga su bazirane na analizi lokalnih karakteristika kao što su srednja vrednost i standardna devijacija intenziteta piksela. Najjednostavniji metod binarizacije je *globalno primena praga* (74) kada se na osnovu histograma odabira jedinstveni prag koji se primenjuje za sve piksele. Ako se prag binarizacije definiše na osnovu lokalnih varijacija intenziteta koristi se *lokalni odabir praga* (75). Najkompleksnija

varijanta lokalne metode odabira praga je dinamičko adaptivni odabir praga kada se prag definiše za svaki pojedinačni i/ili za grupu piksela.

Metode globalne primene praga su pogodne za slike na kojima postoji relativno kontinuiran i izražen kontrast između pozadine i objekta. Takvoj situaciji može da odgovara bimodalni histogram, slika 4.4.2 b. Jednom odabran prag t se primenjuje za sve intenzitete $f(i, j)$ da bi se na kraju dobila binarizovana slika $g(i, j)$ definisana sa

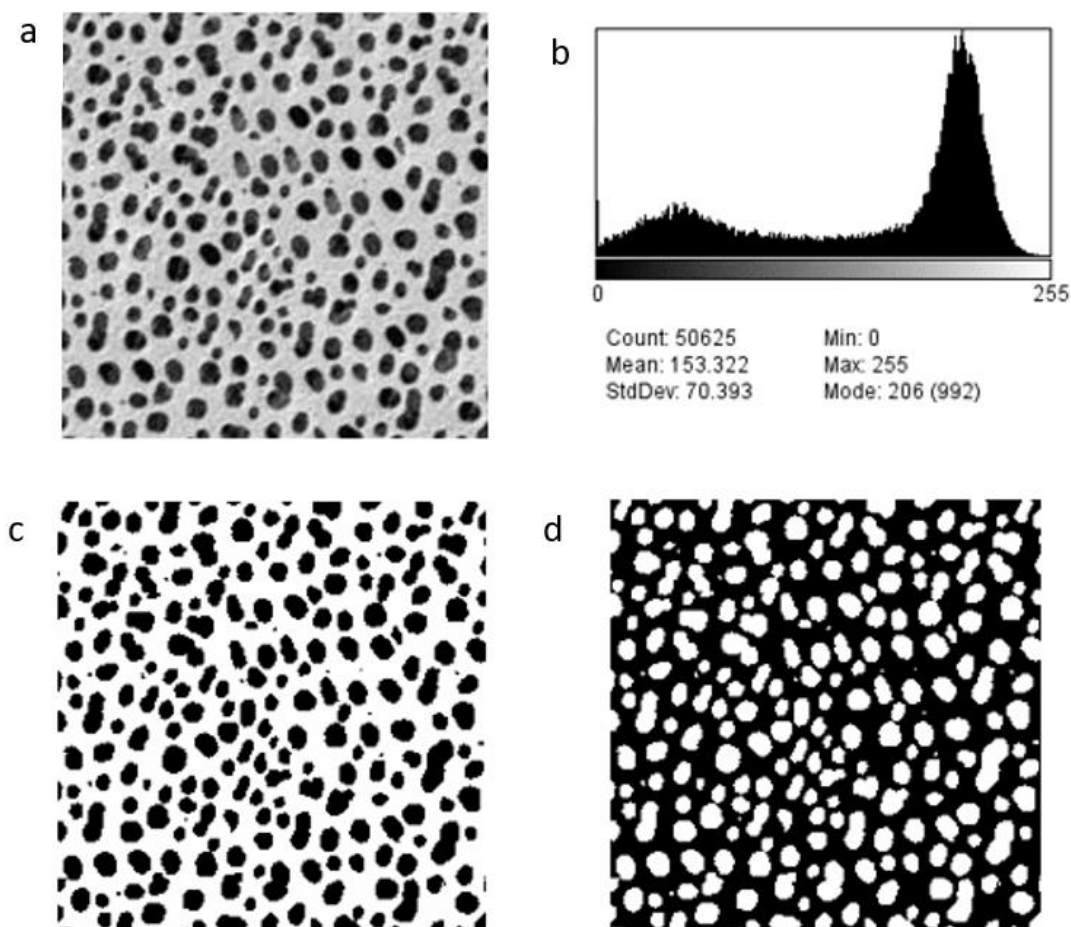
$$g(i, j) = \begin{cases} 1 & \text{za } f(i, j) > t \\ 0 & \text{za } f(i, j) \leq t \end{cases}$$

Globalna primena praga se može realizovati korisnički kada se prema unapred usvojenom kriterijumu odabira određeni intenzitet kao prag ili automatskom primenom neke od metoda koje koriste specifične algoritme za odabir praga. Bez obzira na metod, ideja primene praga je jasna a to je *grupisanje piksela koji pripadaju objektu od interesa i njihovo izdvajanje od pozadine*.

U situacijama kada histogram ne zadovoljava uslove za globalnu primenu praga ili kada globalna primena praga ne dovodi do dobrog razdvajanja objekta od interesa od pozadine primenjuje se lokalno odabran prag. Za određivanje lokalnog praga koriste se algoritmi bazirani na podeli slike na segmente i određivanju praga za svaki segment pojedinačno ili na analizi varijacije intenziteta susednih piksela za svaki pojedinačni piksel. Bez obzira na odabrani metod odabir praga je matematički kompleksan a njegova interpretacije nije tako direktna kao u globalnom odabiru praga.

Metod korisničkog odabira praga je atraktivan jer omogućava korisniku slobodan izbor vrednosti praga ali je skopčan sa problemima reproduktibilnosti, pristrasnosti, izraženom inter- i intra- korisničkom varijabilnosti, vremenom, itd. Te nepogodnosti se na izgled relativno lako rešavaju primenom automatskih metoda odabira praga što takođe može da bude problematično usled postojanja velikog broja algoritama. Među najčešće korišćenim automatskim metodama odabira praga su Huang fuzzy metod, IsoData iterativni metod, metod maksimalne entropije i Otsu. Uspešnost specifične metode izbora praga binarizacije nemoguće je a priori znati.

Uspešnost primene izabrane metode odabira praga je definisana upotrebom vrednošću dobijenih rezultata naknadno izvršenih analiza.



Slika 4.4.2 (a) Originalna 8-bitna slika odabrana kao model i (b) njen histogram. (c) Binarizovana slika sa korisnički odabranim pragom kome odgovara intenzitet $t = 150$. (d) Invertovana binarizovana slika: 0 – pozadina (crno) i 1 – objekti od interesa (belo).

Transformacija slika i varijacija praga binarizacije je urađena korišćenjem programa Fiji/ImageJ verzija 1.53c (71). Prilagođavanje sivih 8-bitnih slika je urađeno globalno primenom korisnički odabranog praga binarizacije i metodom automatske binarizacije. Korišćene su komande: *Process>Binary>Make Binary* za automatsku binarizaciju dok je za korisnički odabir pragova binarizacije korišćeno: *Image>Adjust>Threshold>Set(0,170), (0,200), (0,220), (0,240) i (0,250)* (15). U procesu primene korisnički odabranog praga binarizacije odabrani prag je fiksni i primenjuje se istovetno na svaku pojedinačnu sliku (69).

U procesu automatske binarizacije (*Make Binary*) ImageJ program koristi IsoData iterativni algoritam (76) koji se nezavisno primenjuje na svaku pojedinačnu sliku. U prvoj iteraciji bira se početni prag za koji se, uobičajeno, koristi srednja vrednosti intenziteta piksela koju slika sadrži. Odabranim pragom je izvršena podela histograma na dva dela A i B. Zatim se računaju srednje vrednosti intenziteta za dobijene segmente i_A i i_B koje se koriste za izracunavanje novog praga na sledeći način

$$prag = \frac{i_A + i_B}{2}.$$

Proces se ponavlja sve dok se ne postigne konvergencija rezultata kada je novi prag jednak pragu iz prethodne iteracije.

4.5. Analiza slika i računanje parametara

4.5.1. Analiza teksture digitalnih slika – statistički prilaz

Iako ne postoji opšte prihvaćena definicija teksture očigledno je da je tekstura koncept koji se odnosi na percepciju površine. Baza za upotrebu teksture u opisu i klasifikaciji slika je urođena ljudska sposobnost prepoznavanja osobina površine. U svakodnevnom životu tekstura date slike ili objekta se uglavnom opisuje na kvalitativan način kao fina, gruba, homogena, a skoro nikada se ne uzimaju u obzir kvantitativni parametri. *U analizi digitalnih slika jedna od osnovnih a možda i najprihvaćenija definicija određuje teksturu kao prostornu varijaciju intenziteta piksela koja se kvantifikuje nizom matematičkih parametara.* Prikupljene vrednosti teksturalnih parametara se kasnije koriste pri klasifikaciji slika i identifikaciji objekata, strukturnih elemenata i regiona od interesa (63). U tom smislu tekstura date slike je model koji omogućava kvantifikaciju prostorne varijacije intenziteta piksela koju je uglavnom vrlo komplikovano precizno prepoznati i definisati na neki drugi način (7).

Uzimajući u obzir veliki broj publikovanih radova može se reći da je teksturalna analiza medicinskih slika već tradicionalan metod analize. Međutim, u analizi slika histopatoloških uzoraka nije toliko zastupljena. Još uvek se smatra da je vizuelna

procena histopatoloških uzoraka od strane stručnjaka najmerodavnija pre svega zbog inherentne heterogenosti bioloških tkiva i zbog nestandardizovanih metoda analize (63). Tekstura digitalnih slika histopatoloških uzoraka je posledica osobina tkiva i prostorne raspodele ćelija pri čemu je za očekivati da specifične vrednosti deskriptora teksture imaju biološke i kliničke interpretacije i upotrebne vrednosti (63).

Proučavanju teksture se može pristupiti na više načina. Danas se za te potrebe koristi više aproksimativnih metoda koje se mogu okarakterisati kao metode bazirane na strukturi, na modelima, na transformacijama i statističke metode (77).

Statističke metode su dominantne u analizi teksture biomedicinskih slika. Ove metode dopuštaju opis i analizu površine indirektno upotrebom parametara koji opisuju verovatnoću raspodele intenziteta piksela. U poređenju sa drugim metodama, matematički pristup je opštiji a korišćeni parametri su lakši za računanje i jednostavniji za interpretaciju. Takođe je pokazano i da su ove metode efikasnije i osetljivije u diskriminaciji od ostalih (78).

Tradicionalne statističke metode teksturalne analize se uglavnom baziraju na statistici prvog reda i parametrima dobijenih analizom histograma kojim je predstavljena raspodela učestalosti pojavljivanja spektralnog intenziteta pojedinačnih piksela. Savremenije i matematički kompleksnije statističke metode uzimaju u obzir prostorni kontekst. Polazi se od ideje da kvantifikacija i klasifikacija pojedinačnog piksela zavisi od njegovog okruženja pa time i od načina odabira tog okruženja. Metode koje uzimaju u obzir intenzitet pojedinačnog pikselu i raspodelom intenziteta iz njegovog bližeg ili daljeg okruženja su statističke metode drugog, trećeg i viših redova.

Kombinacija teksturalnih karakteristika sa morfometrijskim deskriptorima može da poveća preciznost histopatološke procene uzoraka u identifikaciji malignih ćelija (79).

STATISTIKA PRVOG REDA – ANALIZA HISTOGRAMA. Karakteristike teksture slike bazirane na statistici prvog reda sadrže informacije koje omogućavaju kvantitativni opis raspodele učestalosti pojavljivanja sivih tonova na osnovu intenziteta piksela. Informacije se predstavljaju u obliku histograma pri čemu se ne uzima u razmatranje

odnos pojedinačnog piksela sa pikselima iz okruženja. Upotreba standardnih veličina, kao što su srednja vrednost ili varijansa, matematički pojednostavljuje primenu statistike prvog reda. Međutim, to je ujedno i jedan od limitirajućih faktora pri diskriminaciji tekstura. Međusobno vrlo različiti slikovni prikazi mogu da rezultuju identičnim vrednostima statistike prvog reda. Uz to je često veoma teško ili gotovo nemoguće povezati korišćene parametre sa karakteristikama uobičajeno korišćenim u kliničkoj praksi.

Za datu teksturu histogram je definisan kao verovatnoća pojavljivanja intenziteta i

$$P(i) = \frac{N(i)}{N} \quad (4.5.1)$$

pri čemu su $N(i)$ broj piksela i -tog intenziteta i N ukupan broj piksela na slici ili u regionu od interesa. Na bazi izraza (4.5.1) se definiše niz statističkih parametara prvog reda koji omogućavaju spektralnu analizu raspodele tonaliteta koji se pojavljuju na slici. Uglavnom se koriste srednja vrednost intenziteta piksela m i centralni momenti μ_k dati sa

$$m = \sum_{i=0}^{N_g-1} i \cdot P(i) \quad i \quad \mu_k = \sum_{i=0}^{N_g-1} (i - m)^k P(i), \quad k = 2,3,4 \quad (4.5.2)$$

gde je sa N_g označen ukupan broj mogućih tonaliteta. Najčešće korišćeni centralni momenti su varijansa (μ_2) koja meri raspršenost rezultata oko srednje vrednosti, koeficijent asimetrije (μ_3) kojim se kvantifikuje forma histograma u odnosu na normalnu raspodelu i koeficijent spljoštenosti (μ_4) koji se koristi kao mera homogenosti raspodele (80).

STATISTIKA DRUGOG REDA – POJAVNE MATRICE (engl. Co-occurrence matrix). Karakteristike teksture slike bazirane na statistici drugog reda uzimaju u obzir odnos između dva piksela, *referentnog* i njegovog *suseda*. Osnovna alatka za izračunavanje statistika drugog reda je pojavna matrica čiji su elementi definisani u funkciji relativne prostorne raspodele piksela. Za izračunavanje elemenata matrice

je neophodna specifikacija prostornog odnosa koji je definisan razdaljinom i uglom između razmatranih piksela (81).

Eksperimenti su ukazali da se diskriminacija tekstura vizuelnim pregledom zasniva na uočavanju razlika u karakteristikama baziranih na statistici kako prvog tako i drugog reda. Za vizuelno razlikovanje slika čija se statistika drugog reda bitno razlikuje nije potreban značajan trud. Međutim, razlikovanje slika koje imaju istovetne ili približne vrednosti karakteristika drugog reda je ili nemoguće ili izuzetno komplikovano (82). Razlike u teksturi slika koje mogu da dovedu do razlika u vrednostima parametara baziranim na statistici trećeg ili višeg reda su verovatno neuhvatljive za ljudsko oko.

Dominantan metod teksturalne analize biomedicinskih slika je GLCM (engl. Grey Level Co-occurrence Matrix) metod kojim se opisuje i analizira frekvencija pojavljivanja kombinacija intenziteta parova piksela u unapred definisanom prostornom odnosu (83). Pokazano je da je primenom GLCM analize moguće doći do informacija koje su povezane sa karakteristikama nevidljivim u vizuelnom mikroskopskom pregledu. Kao na primer, fine promene u strukturi ćelijskog jezgra (84,85) korisne za identifikaciju ćelijskih abnormalnosti, u gradiranju tumora i u dijagnostici (79).

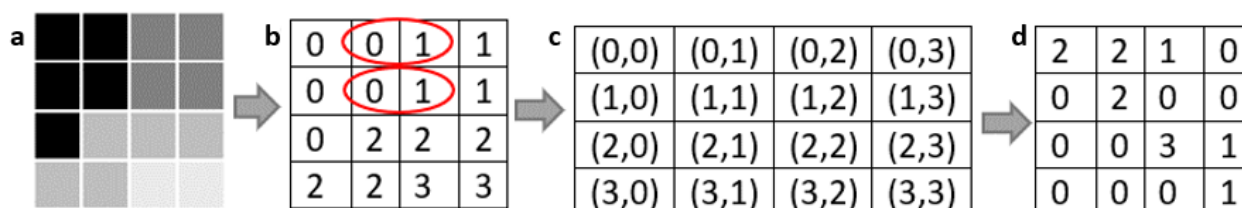
Za izračunavanje elemenata GLCM matrice koristi se uslovna verovatnoća $P(i, j, d, \alpha)$ pojavljivanja i -tog nivo sivog tonaliteta u odnosu na sivi nivo kome odgovara vrednost j a koji se u odnosu na odabrani piksel nalazi na udaljenosti d u pravcu α . Korišćenjem četiri ugla: 0° , 45° , 90° i 135° , dobijaju se četiri matrice koje kada se saberu daju matricu $P(i, j)$ invarijantnu pri rotaciji slike

$$P(i, j) = \sum_{i, j=0}^{N_g-1} P(i, j, d, \alpha) \quad \alpha = 0^\circ, 45^\circ, 90^\circ, 135^\circ. \quad (4.5.3)$$

$P(i, j, d, \alpha)$ je kvadratna matrica stranice jednake broju sivih tonova N_g i kao što je ranije napomenuto može se računati za različite pravce. U opštem slučaju GLCM nije simetrična i osetljiva je na promenu pravca. To se lako prevazilazi uzimajući u obzir resultantnu matricu dobijenu varijacijom uglova. Invarijantna matrica se može normalizovati tako što se vrednost u svakoj ćeliji deli sa ukupnim brojem stanja

$$p(i, j) = \frac{P(i, j)}{\sum_{i, j=0}^{N_g-1} P(i, j)}. \quad (4.5.4)$$

Normalizovana matrica verovatnoća $p(i, j)$ je polazna tačka za izračunavanje vrednosti deskriptora koji se koriste za opisivanje i merenje teksture kao i za klasifikaciju slika.



Slika 4.5.1. Izračunavanje elemenata GLCM matrice. (a) Proizvoljna 2-bitna slika (2^2 tonaliteta su kvantifikovani vrednostima 0, 1, 2 i 3) uzeta za primer. (b) Matrica kvantifikovanih tonaliteta. (c) Nepopunjena i (d) popunjena GLCM matrica.

Izračunavanje elemenata GLCM matrice je višefazni postupak. U prvom koraku se slika zamenjuje matricom intenziteta pridruženim svakom pojedinačnom pikselu (slika 4.5.1. a i b). U sledećem koraku se donosi odluka o prostornom odnosu (d, α) i o pravilu odabira susednog piksela. Na primer, prostorni odnos (1, 0) znači da će se uzeti u obzir prvi desni susedni piksel. Zatim se određuju ulazi (moguće kombinacije datih intenziteta) i popunjavaju ćelije GLCM matrice (slika 4.5.1. c i d). Na primer, u ćeliju (0,1) se unosi broj parova piksela kod kojih referentnom pikselu odgovara intenzitet 0, a prvom desnom susedu, intenzitet 1. Primeru sa slike odgovaraju dve takve kombinacije zaokružene crvenom bojom. Nakon prebrojavanja svih pojavljivanja, definisana je GLCM matrica (slika 4.5.2. d).

Ovako određena matrica nije uzela u obzir krajnju desnu kolonu piksela koji ne poseduju desne susede. Da bi svi pikseli bili uzeti u obzir u analizi, dobijena matrica se zbraja sa transponovanom, pri čemu je dobijena matrica simetrična (slika 4.5.2. a). Transponovanoj matrici odgovara matrica izračunata za prostorni odnos (-1, 0) odnosno za (1, 180°).

U narednom koraku se vrši normalizacija matrice tako što se svaki broj pojavljivanja deli sa brojem mogući pojavljivanja čime se elementi GLCM matrice transformišu u verovatnoće pojavljivanja kombinacija za zadati prostorni odnos (slika 4.5.2 b). Za postizanje potpune uglovne invarijantnosti i uključivanje ivičnih piksela, preporučeni u analizi neorijentisanih slika, uzimaju se u obzir svih osam mogućih uglova prilikom normalizacije.

a	4	2	1	0
	2	4	0	0
	1	0	6	1
	0	0	1	2

b	0.167 (4/24)	0.083 (2/24)	0.042 (1/24)	0 (0/24)
	0.083	0.167 (4/24)	0	0
	0.042	0	0.25 (6/24)	0.042 (1/24)
	0	0	0.042	0.083 (2/24)

Slika 4.5.2. (a) Simetrična matrica dobijena sabiranjem matrice sa slike 4.5.1. d i njene transponovane matrice. (b) Finalna normalizovana matrica verovatnoća pojavljivanja različitih kombinacija sivih tonaliteta za definisani prostorni odnos; u datoj 4×4 matrici postoji 12 kombinacija piksela, tako da je ukupan broj stanja uzet za normalizaciju 24.

Podrazumeva se da su tonalitet i tekstura slike povezani pri čemu u zavisnosti od karakteristika pojedine slike ili jedno ili drugo može da bude dominantno. Velika varijacija u tonalitetu piksela posmatrane oblasti čini da dominantna osobina na datoj površini bude tekstura (83). Uglavnom je prihvaćeno da je kompletan opis teksture sadržan u invarijantnoj GLCM matrici koja je osnova za izračunavanje karakteristika specijalno osmišljenih za njen opis.

Originalni rad (83) definiše 14 parametara teksture od kojih nisu svi međusobno nezavisni i mogu se grupisati prema osobinama koje opisuju. Najčešće korišćeni teksturalni parametri su (86):

- IDM (engl. Inverse Difference Moment). *Inverzni moment razlike* IDM, ili *homogenost*, se računa na sledeći način

$$IDM = \sum_{i,j=0}^{N_g-1} \frac{p(i,j)}{1 + (i - j)^2}. \quad (4.5.5)$$

IDM ima visoke vrednosti kada je matrica skoncentrisana duž dijagonale kao posledica visoke lokalne homogenosti slike. Niske vrednosti kontrasta rezultuju visokim vrednostima IDM.

- ASM (engl. Angular Second Moment). *Ugaoni moment drugog reda* se računa na sledeći način

$$ASM = \sum_{i,j=0}^{N_g-1} p(i,j)^2. \quad (4.5.6)$$

ASM je takođe mera homogenosti ili glatkoće slike. Kada slika poseduje izraženu homogenost i prelazi između tonaliteta nisu dominantni ASM ima visoke vrednosti ASM. Kada su svi pikseli isti ASM ima maksimalnu vrednost (ASM=1). Obrnuto, na slikama velike heterogenosti u tonalitetima ASM ima niske vrednosti.

- CON (engl. Contrast). *Kontrast* ili *inercija* je veličina koja se koristi za opisivanje i merenje heterogenosti slike pomoću merenja lokalnih varijacija. Obrnuto je proporcionalan IDM. Vrednosti CON su visoke kada analizirani region ima visok kontrast. Računa se na sledeći način

$$CON = \sum_{i,j=0}^{N_g-1} p(i,j) \cdot (i-j)^2. \quad (4.5.7)$$

- ENT (engl. Entropy). Približno iste vrednosti elemenata normalizovane matrice generišu visoke vrednosti *entropije*. Kada su vrednosti elemenata matrice bliske 0 ili 1, situacija koja najčešće odgovara uniformnoj slici, vrednosti entropije su niske a slika je glatka. Entropija se računa pomoću

$$ENT = \sum_{i,j=0}^{N_g-1} -p(i,j) \cdot \ln p(i,j) \quad (4.5.8)$$

pri čemu se uzima da je $0 \cdot \ln 0 = 0$. Imajući u vidu da je $p(i,j)$ verovatnoća i kao takva ima vrednosti od 0 do 1, $\ln p(i,j)$ će imati vrednosti manje ili jednake 0. Što je $p(i,j)$ manje to je pojavljivanje date kombinacije piksela ređe a apsolutna

vrednost $\ln p(i, j)$ je veća, i obrnuto. ENT je veličina suprotna ASM i koristi se kao mera heterogenosti slike slično kao i CON.

- COR (engl. Correlation). *Korelacija* ima vrednosti između -1 i 1. Mera je korelacije između piksela u dva pravca i koristi se u kombinaciji sa ostalim parametrima teksture od kojih je uglavnom nezavisna. Korelacija ima veće vrednosti za bliže piksele. Analizirani objekat ima viši stepen korelacije unutar sebe nego sa susednim objektima. COR se računa na sledeći način

$$COR = \sum_{i,j=0}^{N_g-1} \frac{ij \cdot p(i,j) - \mu_i \mu_j}{\sigma_i \cdot \sigma_j} \quad (4.5.9)$$

pri čemu su srednje vrednosti (μ), varijansa (σ^2) i standardne devijacije (σ) određene sa

$$\mu_i = \sum_{i,j=0}^{N_g-1} i \cdot p(i,j), \quad \mu_j = \sum_{i,j=0}^{N_g-1} j \cdot p(i,j)$$

i

$$\sigma_i^2 = \sum_{i,j=0}^{N_g-1} p(i,j) \cdot (i - \mu_i)^2, \quad \sigma_j^2 = \sum_{i,j=0}^{N_g-1} p(i,j) \cdot (j - \mu_j)^2.$$

Standardna devijacije sivih tonaliteta i i j imaće iste vrednosti usled simetrije normalizovane matrice.

Statističke karakteristike trećeg i viših redova uzimaju u obzir prostorni odnos tri i više piksela što prilično usložnjava proračune čak i primenom računara. Pri tome se moraju uzeti u obzir i nedoumice u interpretaciji dobijenih rezultata koji se vrlo teško mogu korelirati sa klinički ustaljenim opisom. Bez potrebe da se ulazi u detalje, neke od najčešće korišćenih metoda su NGLDM (engl. Neighboring Grey Level Dependence Matrix), tj. Q-matrica, i GLRLM (engl. Grey Level Run Length Matrix). Elementi Q-matrice su definisani kao broj pojavljivanja razlike, manje ili jednake a , između piksela i njegovih suseda na razmaku d pri čemu su a i d celi brojevi. Posledica ovakve definicije elemenata je invarijantnost NGLDM na rotaciju

slike i njene nezavisnost od vrednosti ugla (87). Grey Level Run ili texel je niz uzastopnih piksela kojima odgovara isti nivo sivog tonaliteta. Za opis teksture se koristi dužina texela u odabranim pravcima, izražena brojem piksela. Veliki broj uzastopnih piksela istog tonaliteta kojima odgovaraju dugački texeli su karakteristični za isprekidanu teksturu dok su kratki texeli tipični za finu teksturu (87).

4.5.2. Računanje teksturalnih parametara

Za izračunavanje teksturalnih parametara statistike prvog reda i parametara GLCM analize korišćen je program Fiji/ImageJ verzija 1.53c. Parametri statistike prvog reda su izračunate korišćenim celokupnog intervala intenziteta 0 - 255 kao i za svaki od sedam podintervala histograma korišćenjem funkcije *Measure (Analyze>Measure)*. Pre izračunavanja parametara na originalnoj slici je izvršena podela intervala intenziteta odabirom pragova (*Image>Adjust>Threshold*) i prilagođavanjem funkcije *Measure (Analyze>Set Measurements>Limit to threshold)*. Razmatrano je šest parametara: *Površina*, *Srednja Vrednost* intenziteta piksela, *Koeficijenti Asimetrije* i *Spljoštenosti* dati kao treći i četvrti centralni moment definisani formulom 4.5.2, i dve integrisane gustine, *IntDen* data kao proizvod površine i srednje vrednosti intenziteta iz zadatog intervala i *RawIntDen* data kao suma piksela iz zadatog intervala.

Za izračunavanje parametara GLCM analize na 8-bitnim slikama i na sedam setova slika segmentiranih prema prethodno ustanovljenim podintervalima histograma korišćen je podprogram (*Plugins>GLCM Texture*). Podprogram *GLCM Texture* omogućava izračunavanje pet od četrnaest teksturalnih karakteristika (83) najčešće primenjivanih u biološkim i medicinskim istraživanjima (88). GLCM parametri uzeti u obzir su *inverzni moment razlike* (IDM), *angularni drugi moment* (ASM), *kontrast* (CON), *entropija* (ENT) i *korelacija* (COR).

Za izračunavanje *korelacije* je korišćena formula preuzeta iz rada R.F.Walker et al. iz 1995 godine (89)

$$COR = \sum_{i,j=0}^{N_g-1} p(i,j) \frac{(i - \mu_i) \cdot (j - \mu_j)}{\sqrt{\sigma_i \cdot \sigma_j}}. \quad (4.5.10)$$

Korišćena formula je delimično modifikovana u odnosu na originalnu (formula 4.5.9) i najčešće je korišćena formula u modernim softverskim rešenjima koji se bave GLCM analizom.

4.5.3. Savremena morfometrijska analiza histopatoloških slika

Morfometrijska analiza digitalnih slika histoloških uzoraka ima široku primenu i koristi vrlo veliki broj deskriptora, od jednostavnih do izrazito kompleksnih. Nije preterano reći da je kvantitativna histopatologija i citologija postala osnovna komponenta histopatoloških studija. Najznačajnije kliničke primene analize ćelijske i tkivne morfologije su u onkološkim studijama koje parametre koji kvantifikuju morfometrijske karakteristike dovode u vezu sa dijagnozom, predikcijom i prognozom bolesti.

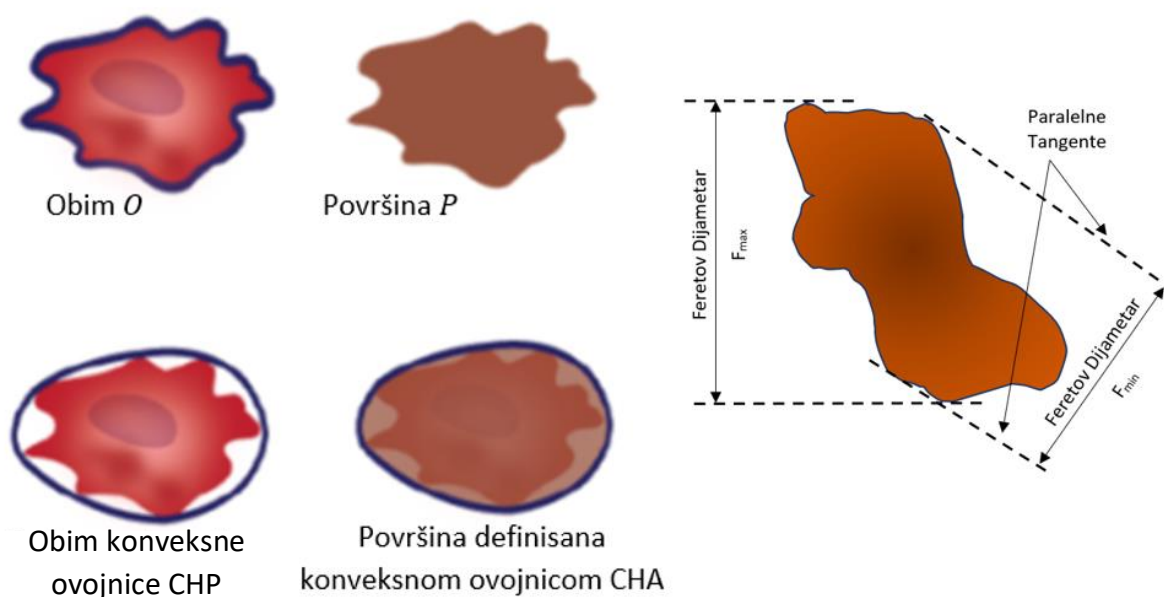
Najjednostavniji deskriptori opisuju osnovne karakteristika ćelijske i tkivne morfologije tako što kvantifikuju veličinu i oblik dvodimenzionalnih bioloških struktura i strukturnih elemenata odnosno njihovih slikovnih prikaza (90) (slika 4.5.3). U slučaju kompleksnih elemenata, što je imajući u vidu diverzitet ćelijskih oblika kod histoloških uzoraka gotovo uvek slučaj, nemoguće je precizno rekonstruisati formu na osnovu specifičnih vrednosti morfometrijskih parametara. Međutim, te vrednosti se uglavnom međusobno dovoljno razlikuju da omogućavaju relativno zadovoljavajuću diskriminaciju.

Kao *deskriptori veličine* najčešće korišćeni parametri su *Obim O* i *Površina P*. Za zatvorenu konturu na slici definisanu kao skup n tačaka sa pripadajućim koordinatama x_i i y_i , obim i površina se određuju na sledeći način (91)

$$O = \sum_{i=1}^{n-1} \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2} \quad i \quad P = \frac{1}{2} \left| \sum_{i=0}^{n-1} (x_i y_{i+1} - x_{i+1} y_i) \right|. \quad (4.5.11)$$

Obim je izračunat kao suma dužina između susednih tačaka granične konture. Površina se odnosi na region obuhvaćen graničnom konturom i na broj konturom obuhvaćenih piksela. Kao alternativa obimu i površini mogu se koristiti obim konveksne ovojnice

CHP (engl.Convex Hull Perimeter) i površina definisana konveksnom ovojnicom CHA (engl.Covex Hull Area) (slika 4.5.3).



Slika 4.5.3. Šematski prikaz morfolometrijskih parametara korišćenih kao kvantifikatori veličine

Morfometrijski parametri najčešće korišćeni kao za kvantifikaciju oblika su:

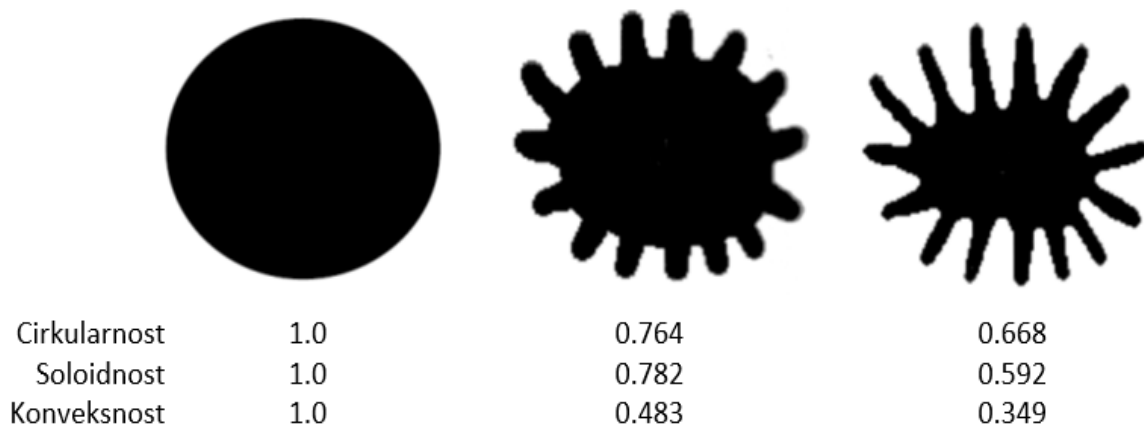
- Cirkularnost (70), takođe poznata i kao kompaktnost (92) ili faktor forme (91,93), definisana je kao odnos između površine i obima pomoću formule

$$Cicircularnost = 4\pi \times \frac{P}{O^2}. \quad (4.5.12)$$

Cirkularnost je mera koincidencije oblika analiziranog objekta i kruga. Objektima kružnog oblika odgovara cirkularnost jednaka 1.0 pri čemu vrednosti manje od 1.0 odgovaraju nepravilnijim objektima i odražavaju kompleksnost granične konture. Što je vrednost cirkularnosti bliža vrednosti 0.0 objekat je izduženiji. Cirkularnost je invarijantna na rotaciju i nezavisna je od veličine objekta. Može se koristiti i kao kvantifikator procesa koji su povezani sa promenom ćelijske forme kao što je promena forme povezana sa apoptozom (94,95).

- *Solidnost* se koristi kao mera konveksnosti i/ili konkavnosti objekta. Vrednost solidnosti jednaka 1.0 odgovara potpuno konveksnim objektima dok su manje vrednosti indikator nepravilnosti graničnog udubljenja odnosno konkavnosti. Solidnost je definisana kao odnos površine ćelije i CHA (96) pomoću formule

$$Solidnost = \frac{P}{CHA} \quad (4.5.13)$$



Slika 4.5.4. Morfometrijski parametri korišćeni za kvantifikaciju geometrijskih karakteristika objekata različitog oblika.

Kao kvantifikator oblika i veličine često se koristi i Feret-ov dijametar definisan kao maksimalno rastojanje između dvaju paralelnih tangenti F_{max} (slika 4.5.3). Takođe se koriste i minimalni Feret-ov dijametar, *FeretAngle* definisan kao ugao između Feret-ovog dijametra i x-ose, *FeretX* i *FeretY* koje odgovaraju početnim koordinatama F_{max} .

Za kvantitativni opis oblika ćelija i ćelijskih elemenata često se upotrebljavaju i:

- *zaobljenost* (engl. roundness) definisana kao $1/\text{cirkularnost}$ (97) tako da savršeni krug ima vrednost 1.0 a vrednostima većim od 1.0 su opisani izduženi oblici,
- *koeficijent aspekta* (engl. aspect ratio) je definisan kao odnos najveće i najmanje ose koje predstavljaju najveće i najmanje rastojanje između dve tačke granične konture (70) tako da krugu odgovora vrednost 1.0,

- *ekscentričnost* (98) je definisana kao odnos najveće i najmanje ose najbolje prilagođene elipse koja sadrži analizirani objekat. Vrednost 0,0 odgovara savršenom krugu a vrednost 1.0 pravougaoniku.
- *eliptičnost* je definisana kao odnos cirkularnosti i cirkularnosti najbolje prilagođene elipse (99) tako da eliptičnoj formi odgovara vrednost 1.0,
- *konveksnost* je definisana kao odnos obima konveksne ovojnice i obima objekta (100) i mera je gubitka konveksnosti.

Usled nedostatka standardizacije i izražene heterogenosti tkiva i ćelija, pored navedenih parametara koriste se i mnogi drugi morfometrijski kvantifikatori. Na primer, za opis složenih neuronskih oblika se koriste broj tačaka granjanja, gustina granjanja, faktor ramifikacije, radijalna udaljenosti, ugao granjanja (101,102) kao i parametri fraktalne geometrije (103).

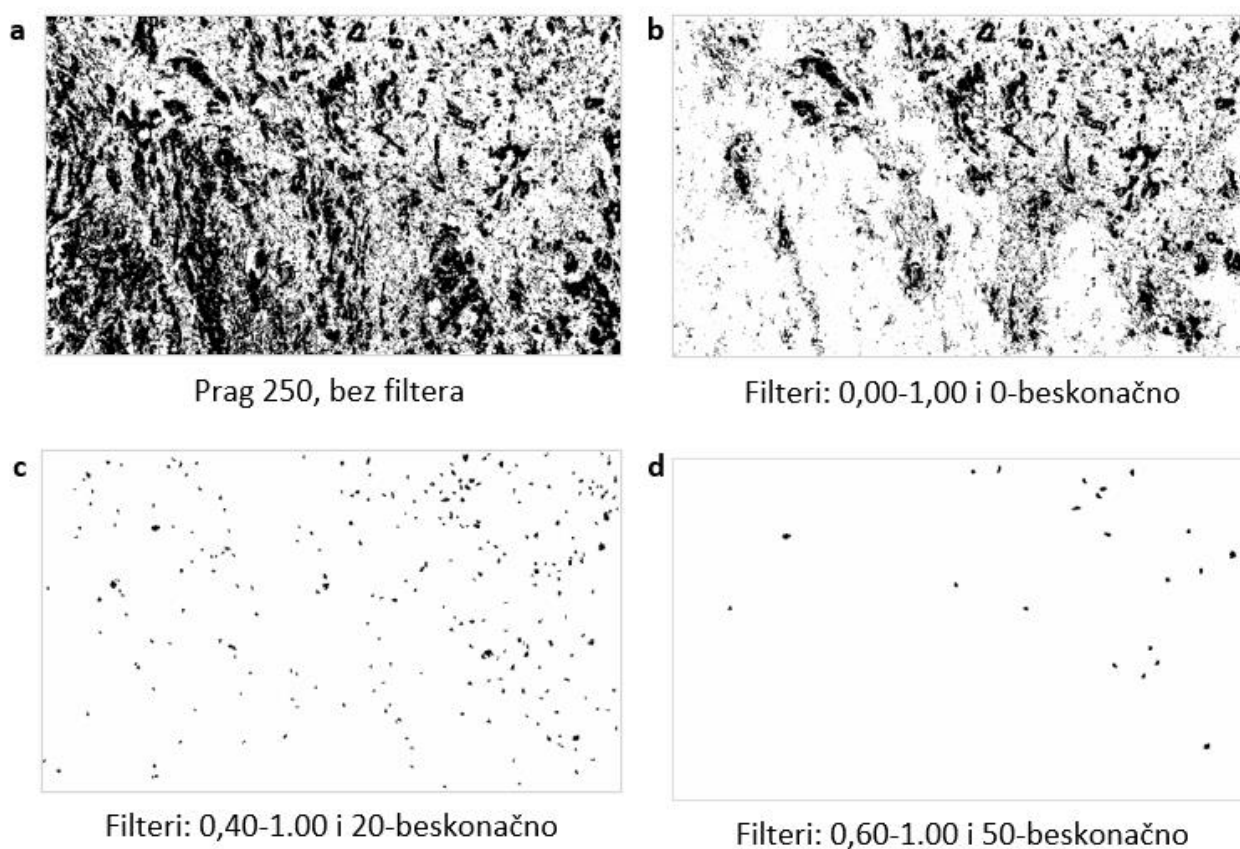
Analiza morfologija zasnovana na morfometrijskim karakteristikama digitalnih slika histopatoloških uzoraka susreće sa poteškoćama koje se mogu grupisati na probleme povezane sa pripremom uzorka i probleme povezane sa analizom slike. Za uspešnu analizu je neophodna priprema uzorka koja uključuje ekstrakciju, fiksaciju i bojenje (50).

Iako je porast u broju algoritama korišćenih u računarskoj morfometrijskoj analizi impozantan još uvek postoje problemi koji zahtevaju efikasno rešenje. Za preciznu i pouzdanu analizu pre svega su neophodni algoritmi koji na konzistentan način omogućavaju identifikaciju i segmentaciju objekata od interesa što je preduslov za nepristrasno računanje morfometrijskih parametara (90).

Izražena varijabilnost u protokolima prikupljanja slika kao i inherentna varijabilnost biološko-medicinskih uzoraka zahteva algoritme koji mogu da se nose na pouzdan i precizan način sa raznolikošću tipičnom za digitalne slike histopatoloških uzoraka.

4.5.4. Računanje broja i morfometrijskih parametara epitelnih malignih plaža

U proseku je za svakog od 102 pacijenta analizirano pet slika. Optimizacijom praga binarizacije je dobijeno šest setova slika koje su podvrgnute morfometrijskoj analizi korišćenjem programa ImageJ i podprograma *Analyze Particles*⁶ (*Analyze>Analyze Particles*) koji automatski vrši segmentaciju slike izdvajajući elemente predefinisanih karakteristika (veliĉine i oblika) od pozadine.



Slika 4.5.5. Efekat filtriranja veliĉine i oblika ćelijskih plaža. (a) Binarizovana slika sa pragom podešenim na 250. (b) Identifikovane i ocrtane plaže bez primene filtera. (c) Plaže identifikovane primenom filtera 0,40 - 1,00 (cirkularnost) i 20-beskonaĉno (veliĉina) i (d) Rezultat primene filtera 0,60 - 1,0 i 50 - beskonaĉno.

⁶ Korišćeni podprogram za analizu klastera epitelnih ćelija koristi termin *ĉestica* (engl. particle). Shodno tome termin *ĉestica* će se koristiti kao sinonim termina klaster, plaža i izdanak tumora, a parametar kome odgovora broj identifikovanih obojenih pojedinaĉnih klastera na slici, tj. broj malignih plaža i izdanaka tumora, je nazvan *Broj ĉestica*.

Svaki pojedinačni set binarizovanih slika je segmentovan u funkciji veličine i oblika obojenih epitelnih malignih plaža (slika 4.5.5. b - d) čija je filtracija izvršena prema dvema karakteristikama: cirkularnost i veličina. Cirkularnost čije vrednosti idu od 0,0 do 1,0 je menjana u intervalima 0,0 - 1,0, 0,2 - 1,0, 0,4 - 1,0, 0,6 - 1,0 i 0,8 - 1,0. Maksimalna vrednost filtera veličine je fiksirana na beskonačnosti dok su za minimum uzimane vrednosti 10, 20, 50 i 100. Kombinacija odabranih filtera (prag binarizacije šest, cirkularnost pet i veličina četiri filtera) je rezultovala generisanjem 120 setova slika.

Filtriranjem veličine i cirkularnosti je moguće identifikovati ćelijske plaže u funkciji veličine i oblika. Efekat filtriranja na broj identifikovanih plaža, prosečnu veličinu, cirkularnost, solidnost i Feretov dijametar se može posmatrati na primeru datim na slici 4.5.5. Slika uzorka sa pragom binarizacije postavljenim na 250 je prikazana na slici 4.5.5. a. Na slici b su prikazane plaže identifikovane bez primene filtera tako što su uzeti u obzir svi ćelijski klasteri. Izbrojano je 11526 plaža. Ignorišući plaže veličine manje od 20 piksela i ograničavajući cirkularnost na vrednosti veće od 0,40 (slika 4.5.5. c) broj plaža je smanjen na samo 252. Daljim povećanjem pragova filtracije na 0,60 i 50 efekat je bio ekstremno, identifikovana je samo 21 plaža (slika 4.5.5. d).

Za svaku pojedinačnu sliku iz svakog pojedinačnog seta klasteri su prebrojani i analizirana su njihove morfometrijske karakteristike. Među analiziranim parametrima se nalaze *Broj čestica*, parametri veličine (*Ukupna površina* svih čestica i njihova *Prosečna veličina* mereni u pikselima, obim), parametri oblika (*Cirkularnost*, *Feretov dijametar* i *MinFeret*) i parametri gustine (*Solidnost*).

4.6. Procena učinka prognostičkih faktora u medicini

U modernoj medicinskoj praksi je ustaljena upotreba modela koji uključuju testove, diskriminatore, klasifikatore i kriterijume u dijagnostičke, prediktivne i prognostičke svrhe. Da bi model bio prihvaćen efikasnost predviđanja mora da zadovolji kriterijume preciznosti i generalizacije (104). Potvrda preciznosti predviđanja modela se proverava u procesu interne validacije a jedan od uslova koji se mora zadovoljiti je prihvatljiv nivo *diskriminacije*. Za prognostičke modele je

neophodno da uspešno obavljaju diskriminaciju pacijenata u odnosu na odabrani ishod bolesti. U tom cilju je neophodno izvršiti procenu učinka modela poređenjem rezultata dobijenih testiranjem sa retrospektivno utvrđenim ishodom bolesti (105).

Po svojoj nameni prognostički faktori su klasifikatori koji se primenjuju za klasifikaciju pacijenata prema unapred prihvaćenom kriterijumu. Po svojoj prirodi, prognostički faktori su parametri koji u zavisnosti od karakteristike koju kvantifikuju mogu biti kontinuirane ili diskretne veličine. Kada se za diskriminaciju koriste kontinuirani parametri potrebno je ustanoviti prag, tj. graničnu vrednost koja određuje granicu klase. U procesu klasifikacije vrednost prognostičkog parametra se poredi sa vrednošću praga što rezultuje svrstavanjem subjekta u jednu od klasa (106).

U medicini je uobičajeno da se za kvantifikaciju efikasnosti diskriminacije dihotomnih parametara koriste dve veličine: *Senzitivnost* i *Specifičnost*. Obe ove veličine su po svojoj prirodi uslovne verovatnoće (107). *Senzitivnost* (*SPS*) je verovatnoća za dobijanje pozitivnog rezultata u testiranju bolesnih osoba dok je *Specifičnost* (*SNS*) verovatnoća dobijanja negativnih rezultata u testiranju zdravih osoba (108). Na primer, ako je od 100 bolesnih osoba njih 62 identifikovano kao pozitivno tada je *Senzitivnost* primenjenog testa 62%. *Senzitivnost* je definisana kao odnos tačno pozitivnih (*TP*) osoba i ukupnog broja pozitivnih osoba (*P*). Pri tome je *TP* broj ispravno identifikovanih bolesnih osoba a *P* čine *TP* pacijenti i lažno negativne (*FN*) osobe, tj. pogrešno identifikovani bolesne osobe. Na sličan način se definiše i *Specifičnost* kao odnos tačno negativnih (*TN*) i ukupnog broja negativnih osoba (*N*). Pri tome je *TN* broj ispravno detektovanih zdravih osoba a *N* čine *TN* pacijenti i lažno pozitivne (*FP*) osobe, tj. pogrešno identifikovane zdrave osobe. Odgovarajući izrazi za *Senzitivnost* i *Specifičnost* su

$$SPS = P(TP|P) = \frac{TP}{TP + FN} \quad i \quad SNS = P(TN|N) = \frac{TN}{TN + FP} \quad (4.6.1)$$

pri čemu su $P = TP + FN$ i $N = TN + FP$. Na osnovu definicije proizilazi da su *Senzitivnosti* i *Specifičnosti* inherentna svojstva testa koja ne zavise od prevalencije analiziranog stanja. Alternativni parametri za *Specifičnost* i *Senzitivnost* su lažno

pozitivna stopa (FPS) i lažno negativna stopa (FNS). Pri tome su $FPS = 1 - SNS$ i $FNS = 1 - SPS$.

Na osnovu formule (4.6.1) *Senzitivnost* se može razumeti i kao mera izbegavanja FN pacijenata. Stoga je negativan rezultat u testu visoke *Senzitivnosti* prihvatljiv za odbacivanje bolesti. Obrnuto ne važi. Naime, test za koji je $SPS = 1$ i čija je *Senzitivnost* 100% će sve bolesne subjekte prepoznati kao pozitivne. U tom slučaju negativan rezultat definitivno isključuje postojanje bolesti. Međutim, pozitivan rezultat nije nužno pogodan za diskriminaciju. *Senzitivnost* ne uzima u obzir pogrešno dijagnostikovane zdrave pacijente tako da i u testu koji uopšte ne vrši diskriminaciju SPS može biti maksimalna sa 100%. Identična logika je primenjiva na tumačenje *Specifičnosti* kao mere izbegavanja FP pacijenata.

Može se zaključiti da u idealnom slučaju savršen diskriminator ima i SPS i SNS od 100%. To u praksi nikad nije slučaj. Da bi diskriminator imao dobru upotrebnu vrednost potrebno je da se obezbedi optimalan kompromis između SPS i SNS . Optimalna vrednost *Senzitivnosti* i *Specifičnosti* zavisi od kliničke upotrebe testa (109). Ozbiljne bolesti zahtevaju senzitivnije testove. Na primer, visoko senzitivni test je primeren u visoko rizičnoj grupi za rak čak i kada ima visoku stopu FP pacijenata jer se smanjuje verovatnoća FN rezultata što je u slučaju raka prihvatljivije.

Za procenu učinka diskriminatora pri fiksiranoj vrednosti praga klasifikacije moguće je koristiti i *Tačnost* (ACC) koja se definiše kao verovatnoća dobijanja ispravnog rezultata

$$ACC = \frac{TP + TN}{P + N}. \quad (4.6.2)$$

No kako je *Tačnost* zavisna i od SNS i od SPS , naročito kada nisu izbalansirane i kada je ACC bliža većoj veličini (80), preostaje kao najbolji kriterijum procena *Senzitivnosti* i *Specifičnosti*. Uobičajeno je da se procena optimalnih vrednosti *Senzitivnosti* i *Specifičnosti* pri variranju praga diskriminacije vrši metodom poznatom kao *ROC analiza* (106).

Analiza nezavisnosti i uticaja pojedinačnih prognostičkih faktora na vreme preživljavanja najčešće se ispituje primenom multivarijacione Cox-ove regresione

analize. Cox-ov regresioni model proporcionalnog rizika je semi-parametarski model koji omogućava istovremenu analizu efekata više numeričkih i/ili kategoričkih faktora rizika na vreme preživljavanja kao i kvantifikaciju i poređenje rizika između grupa pacijenata. Realističniji je od Kaplan-Meier modela jer predviđa da rizik varira u toku vremena ali na takav način da je odnos rizika *HR* (engl. Hazard Ratio) između grupa konstantan. Konstantan *HR* omogućava kvantitativnu procenu efekta bez poznavanja funkcija osnovnog rizika. Za razliku od ROC analize Cox-ov model uzima u obzir faktor vremena, međutim, zahteva kategorizaciju vrednosti korišćenih parametara na osnovu arbitrarno usvojenog praga što mu je ujedno i jedan od glavnih nedostataka (15).

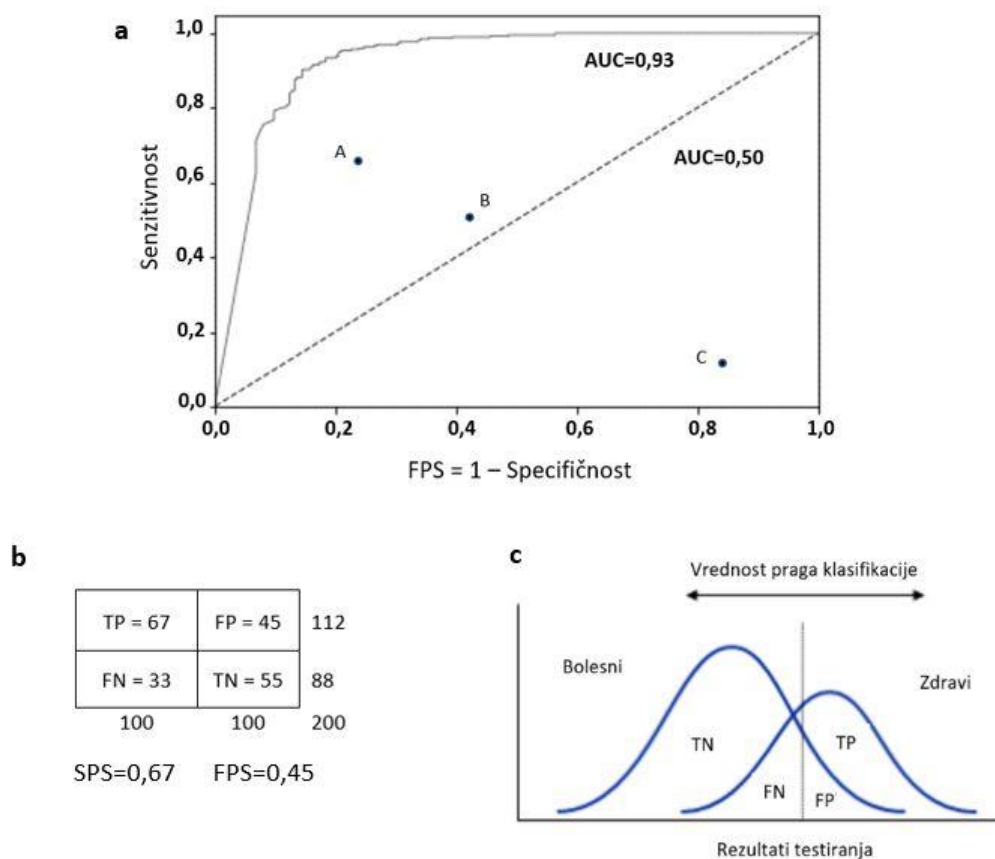
4.6.1. ROC analiza u medicinskim istraživanjima

U medicinskim istraživanjima za procenu diskriminatornog učinka kontinuiranih parametara pri različitim vrednostima praga klasifikacije uglavnom se koriste ROC krive i ROC analiza. ROC krive su grafici *SPS* prema *FPS*, tj. *Senzitivnost* prema *1-Specifičnosti*, dobijeni variranjem praga diskriminacije čime je omogućena jednostavna vizualizacija inverznog dinamičkog odnosa *SPS* i *SNS* (slika 4.6.1. a).

Rezultati ROC analize imaju višestruku upotrebnu vrednost (109). U kontekstu prognoze se koriste za procenu učinka prognostičkih parametara odnosno njihove diskriminatorne efikasnosti u klasifikaciji subjekata. Kao kvantitativna mera ostvarenog diskriminatornog učinka se koristi AUC, to jest površina ispod ROC krive. AUC vrednost odgovara verovatnoći da će primenjeni klasifikator od dva slučajno odabrana pacijenta od kojih je jedan bolestan (pozitivan) a drugi zdrav (negativan) rangirati pozitivnog kao višeg u uz pretpostavku da pozitivnim pacijentima odgovara viši rang nego negativnim (106).

ROC analiza se koristi i za ustanovljavanje optimalne vrednosti praga klasifikacije kojim se postiže kompromis između *Senzitivnosti* i *Specifičnosti*. Uobičajeno je i da se koristi za poređenje testova (110) kao i za analizu posmatračke varijabilnosti.

KONSTRUKCIJA ROC KRIVE I TUMAČENJE REZULTATA. Pri konstrukciji ROC krive (slika 4.6.1. a) na y -osu se nanose vrednosti *Senzitivnosti* a na x -osu vrednosti koje odgovaraju $1 - \textit{Specifičnost}$ (106). U tako definisanom koordinatnom sistemu *Senzitivnost* raste u pravcu porasta vrednosti y -ose. Za *Specifičnost* važi obrnuto. Što je vrednost x -koordinate bliža nultoj vrednosti *Specifičnost* je veća i dostiže svoju maksimalnu vrednost $SNS = 1$ pri $FPS = 0$.



Slika 4.6.1. (a) ROC krive i AUC vrednosti različitih diskriminatora. (b) Tabla kontigencije za računanje pojedinačnih vrednosti *Senzitivnosti* i *Specifičnosti*. (c) Uobičajeno preklapanja u raspodeli klasifikovanih pacijenata.

Uzimajući u obzir da se vrednosti SPS i FPS menjaju u intervalu od 0 do 1, uobičajeno je da se ROC kriva konstruiše u ROC prostoru koji čini kvadrat stranice dužine 1. Tačke koje leže na ROC krivi su određene parovima vrednosti SPS i FPS izračunatim variranjem praga diskriminacije pri čemu je opšte prihvaćeno da se pojedinačni rezultati prikazuju u obliku matrica kontigencije (slika 4.6.1. b).

Dijagonala koja spaja naspramna temena kvadrata sa koordinatama (0, 0) i (1, 1) deli ROC prostor na dva identična dela površine 0,5. Diskriminatoru čija ROC

kriva odgovara dijagonali se pridružuju $AUC = 0,5$ i $ACC = 50\%$ te stoga taj diskriminator nije prihvatljiv jer razvrstava pacijente potpuno nasumično (111). Krećući se duž dijagonale *Senzitivnost* raste a *Specifičnost* opada tako da su pojedinačnim rezultatima koji "padaju" na dijagonalu šanse za dobijanje ispravnog rezultata 50-50.

Tački (0, 0) u ROC prostoru odgovaraju maksimalna *Specifičnost* i minimalna *Senzitivnost*, $SPS = 100\%$ i $SNS = 0\%$. Pripadajući prag klasifikacije očigledno nema upotrebnu vrednost. Situacija koja odgovara naspramnom temenu u tački (1, 1) je obrnuta. U toj tački je $SPS = 0\%$ a $SNS = 100\%$. Usled nepostojanja kompromisa između *Specifičnosti* i *Senzitivnosti* pripadajući prag klasifikacije, kao i u prethodnom slučaju, nije prihvatljiv. Napredujući duž *x*-ose, u pravcu rasta *FPR*-vrednosti, vrednosti *Senzitivnosti* su fiksirane na nulu dok *Specifičnost* opada tako da u tački (1, 0) dolazi do "najneprihvatljivijeg" rezultata za koji su i $SPS = 0$ i $SNS = 0$. Suprotnom temenu u tački (0, 1) odgovara najbolji mogući rezultat jer se u toj tački postiže maksimalna vrednost i *Senzitivnosti* i *Specifičnosti* sa $SPS = 100\%$ i $SNS = 100\%$. Prag klasifikacije koji odgovara ovoj tački doveo bi do idealne klasifikacije pacijenata.

Ekstrapolacijom prethodne analize na preostale tačke u ROC prostoru dolazi se do zaključka da tačkama iznad dijagonale u gornjoj polovini prostora odgovaraju bolji rezultati pri čemu veća udaljenost tačke od dijagonale ukazuje na veću efikasnost klasifikacije. Najudaljenijoj tački odgovara položaj (0, 1). Kriterijum udaljenosti tačke od dijagonale je primenljiv i na tačke raspoređene ispod dijagonala u donjoj polovini ROC prostora.

Tako "najneprihvatljivijoj" tački, temenu (0, 1), u stvari odgovara idealna diskriminacija pri čemu je klasifikacija invertovana te su negativno klasifikovani pacijenti bolesni a pozitivni pacijenti zdravi. Inverzija rezultata klasifikacije, negativno u pozitivno i obrnuto, dovodi do rotacije ROC krive oko dijagonale u deo prostora kome odgovara ispravna klasifikacija pacijenata.

Shodno usvojenom tumačenju, vizuelnim pregledom primera sa slike 4.6.1. a se zaključuje da tački C u invertovanoj situaciji odgovara najbolji a tački B najlošiji rezultat.

Proširenjem prethodne analize na različite raspodele tačaka u ROC prostoru i različite krive dolazi se i do tumačenja pripadajućih AUC vrednosti koje mogu imati vrednosti od 0,0 do 1,0 i koje na numerički način opisuju efikasnost diskriminacije. Dijagonali odgovara $AUC = 0,5$ i shodno prethodnoj analizi ova vrednost i odgovarajuća kriva nemaju upotrebnog značaj jer su rezultat nasumične klasifikacije. Maksimalna vrednost površine koja se može postići je $AUC = 1,0$. Ova vrednost odgovara idealnom slučaju (rezultat savršene diskriminacije) (111), tj. situaciji u kojoj nema preklapanja u raspodeli pozitivnih i negativnih subjekata (slika 4.6.1. c). Ovaj slučaj u praksi je nemoguće postići. AUC vrednosti veće od 0,5 su u negativnoj korelaciji sa prognozom tako da AUC vrednosti iz intervala 0,5 - 1,0 ukazuju na povezanost sa visokim rizikom za odabrani ishod, i obrnuto, AUC vrednosti iz intervala 0,0 - 0,5 ukazuju na povezanost sa niskim rizikom od ishoda (61). Kao što je ranije napomenuto, u analizama izvedenim za potrebe ove teze kao ishod od interesa je prihvaćena pojava metastaza.

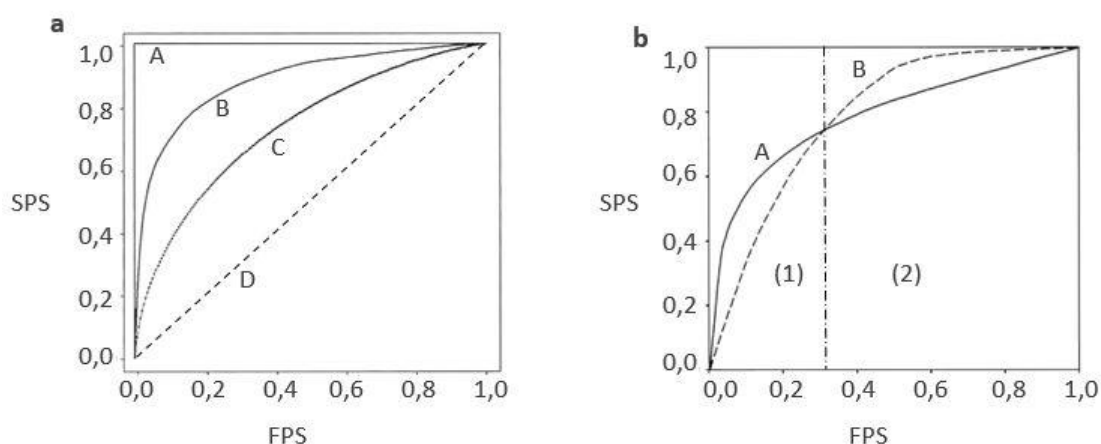
Za kvalitativnu procenu prognostičkog učinka je uobičajeno da se intervali AUC vrednosti od 0,5 do 1,0 i od 0,5 do 0,0 izdele na podintervale. U zavisnosti u kom podintervalu se nalazi postignuta AUC vrednost učinak analiziranog parametra se može kvalifikovati kao dobar, vrlo dobar, itd. Za potrebe ove teze je prihvaćena sledeća raspodela AUC vrednosti i sledeća procena učinka: a) između [0,50, 0,60) učinak je *nedovoljno dobar*, b) između [0,60, 0,70) učinak je *dobar*, c) između [0,70, 0,80) učinak je *vrlo dobar*, d) između [0,80, 0,90) učinak je *odličan* i e) između [0,90, 1,00] učinak je *skoro savršen*.

Parametri čije se AUC vrednosti nalaze u intervalu 0,4 - 0,6 su bez prognostičkog značaja usled potpunog odsustva ili vrlo loše diskriminacije. AUC vrednosti izvan ovih intervala (veće ili jednake 0,6, manje ili jednake 0,4) su *prognostički značajne AUC vrednosti*. *Prognostički značajni parametri* postižu prognostički značajne AUC vrednosti i sposobni su da vrše klasifikaciju pacijenata na prihvatljiv način.

AUC vrednosti manje od 0,5 koje se nalaze ispod dijagonale nisu obavezno loše. Imajući u vidu inverziju klasifikacije, interpretacija "negativnih" vrednosti je analogna interpretaciji "pozitivnih" vrednosti pa je tako $AUC = 0,5 - 0,4$ nedovoljno

dobar rezultat, 0,4 – 0,3 dobar, 0,3 – 0,2 je vrlo dobar, 0,2 – 0,1 je odličan a 0,1 – 0,00 skoro savršen rezultat pri čemu $AUC = 0,0$ odgovora idealnoj klasifikaciji.

Prihvatanjem AUC vrednosti kao indeksa u oceni performansi prognostičkih parametara omogućeno je direktno poređenje testova: što je AUC vrednost veća prognostička moć testa ja prihvatljivija (112). Na slici 4.6.2 a, parametar kome odgovara kriva A je idealan, parametar kome odgovara kriva B ima bolji prognostički učinak od parametra kome odgovara kriva C dok kriva D odgovara nasumičnoj klasifikaciji.



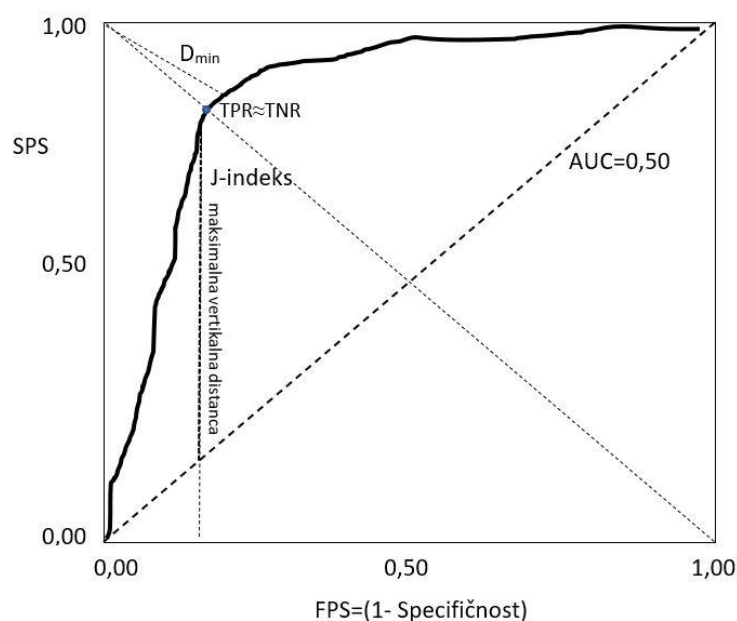
Slika 4.6.2. (a) Poređenje testova u ROC prostoru: krivoj A odgovara idealna podela ($AUC = 1,0$) dok kriva D ne vrši diskriminaciju ($AUC = 0,5$). Takođe je $AUC_B > AUC_C$. (b) Različiti testovi jednakih AUC vrednosti.

Vredi napomenuti da dve različite krive mogu da imaju istovetnu AUC vrednost (110) što se događa kada se krive seku (slika 4.6.2. b). U tom slučaju, ako kliničke okolnosti zahtevaju korišćenje testa veće specifičnosti, deo (1) na slici, preporučuje se korišćenje testa A, dok je u drugom delu ROC prostora, delu (2), test B specifičniji od testa A.

ODREĐIVANJE PRAGA DISKRIMINACIJE. Za efektivnu kliničku primenu prognostičkog modela neophodno je određivanje optimalnog praga klasifikacije. U zavisnosti od specifičnih i konkretnih kliničkih okolnosti i potreba, u izvesnim situacijama je potrebno obezbediti veću senzitivnost dok je u nekim drugim situacijama poželjna veća specifičnost. U situacijama kada nije izražena

preferencijalnost između ovih veličina teži se postizanju njihovih maksimalnih vrednosti tako da su modeli određivanja praga bazirani na postizanju optimalnog kompromisa između *Senzitivnosti* i *Specifičnosti* (111). Najčešće primenjivani kriterijumi određivanja praga klasifikacije se baziraju na upotrebi:

- 1) *Youden indeks*, $J = \max[SPS - (1 - SNS)]$. Ovim indeksom je zapravo izražena težnja postizanja maksimalne razlike između *SPS* i *FPS* tako da J-indeksu grafički odgovara maksimalna vertikalna udaljenost u odnosu na dijagonalu (slika 4.6.3.). J-indeks je jednostavan za izračunavanje i među najčešće korišćenima.



Slika 4.6.3. Određivanje optimalnog praga za klasifikaciju.

- 2) D_{\min} , minimalno rastojanje između AUC krive i temena (1, 0). Identifikuju se koordinate, (slika 4.6.3.) kojima odgovara minimalna vrednost

$$D = \sqrt{(1 - SPS)^2 + (1 - SNS)^2}.$$

Ova tačka je najbliža idealnoj klasifikaciji kojoj odgovara $SPS = 1$ i $SNS = 1$.

- 3) $SPS \approx SNS$, identifikaciju se tačke u kojoj su vrednosti *Senzitivnosti* i *Specifičnosti* najpribližnije jedna drugoj. Ta tačka se nalazi u preseku ROC krive i dijagonale koja spaja temena (0, 1) i (1, 0) i za nju površina $SPS \times SNS$ dostiže maksimalnu vrednost.

Prednosti korišćena ROC-AUC analize su višestruke: ROC krive su nezavisne od prevalencije, AUC je invarijantna pri različitim transformacijama (npr. linearizacija ulaznih varijabli), analiza se vrši za sve moguće vrednosti praga što direktno omogućava procenu optimalne granične vrednosti klasifikacije i omogućeno je direktno poređenje učinka testova (109). Određeni faktori mogu da dovedu do pristrasnosti u tumačenju rezultata i često nije moguće da se kompletno eliminišu iz procesa. Među ometajućim faktorima za izvođenje ROC analize nalaze se tačnost retrospektivnog ishoda, komorbiditet, interobservabilnost, itd.

4.6.2. Cox-ov regresioni model proporcionalnog rizika

U analizi preživljavanja se primenjuju niz statističkih metoda za koje je promenljiva od interesa *vreme do pojave ishoda*. Ishod od interesa se naziva i neuspeh usled toga što je najčešće kao ishod odabran negativan događaj, na primer, pojava bolesti, relapsa, smrt, i slično. Ishod može biti i pozitivan, na primer, oporavak, povratak na posao, odgovor na tretman, i slično.

Za opis raspodele vremena preživljavanja uglavnom se koriste dve funkcije, *funkcija preživljavanja* $S(t)$ i *funkcija rizika* $h(t)$

$$S(t) = P(T > t) \quad i \quad h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t < T < t + \Delta t | T \geq t)}{\Delta t}. \quad (4.6.3)$$

Funkcija preživljavanja je verovatnoća da pacijent preživi duže od vremena t . Funkcija rizika je definisana kao uslovna verovatnoća da pacijent doživi ishod neposredno posle trenutka t ukoliko je preživeo do trenutka t i predstavlja trenutnu stopu smrtnosti (113). Na osnovu definicije se zaključuje da je funkcija preživljavanja monotona ne rastuća funkcija vremena t za koju je $S(t = 0) = 1$ i $S(t = \infty) = 0$. Teorijski kriva funkcije preživljavanja je glatka, međutim u realnim istraživanjem funkciji $S(t)$ odgovara stepenasta kriva. Funkcija preživljavanja, poznata i kao kumulativna stopa preživljavanja, se uglavnom koristi za direktno vizuelno poređenje obrazaca preživljavanja u grupama, procenu hazarda, srednjeg vremena preživljavanja, između ostalog. Za razliku od funkcije preživljavanja, funkcija rizika može rasti, opadati, ili biti konstantna.

Odnos ove dve funkcije je inverzan. To je i logično jer je funkcija preživljavanja fokusirana na preživljavanja dok je funkcija rizika fokusirana na pojavu ishoda tako da dok jedna raste druga opada i obrnuto. Funkcije preživljavanja i rizika su međusobno ekvivalentne. Ako je poznata jedna može se odrediti druga korišćenjem jedne od relacija

$$h(t) = -\frac{d}{dt} \log S(t) \quad \text{tj.} \quad S(t) = e^{-\int_0^t h(u) du}. \quad (4.6.4)$$

Analizom podataka uzorka se procenjuje jedna od ove dve funkcije što omogućava donošenje zaključaka o obrascu preživljavanja.

U analizama preživljavanja najčešće su korišćeni (113):

- 1) *Kaplan-Meier-ov (KM) model*; KM model je ne-parametarski model jednostavan za interpretaciju koji na osnovu vremena do ishoda omogućava procenu verovatnoće preživljavanja (114). Funkciji preživljavanja u ovom modelu odgovara stepenasta funkcija čija se vrednost menja svaki put kad neki od pacijenata doživi ishod. Kaplan-Meier-ove krive preživljavanja omogućavaju procenu intervala pouzdanosti verovatnoće preživljavanja, poređenje preživljavanja u grupama i procenu srednjeg vremena preživljavanja. Međutim, ovaj metod opisuje preživljavanje u funkciji samo jedne ili nekoliko kategoričkih veličina dok se uticaj ostalih zanemaruje. Takođe, ne postoji jednostavna matematička funkcija koja opisuje krive preživljavanja i nije moguće izračunavanje stope rizika kao kvantifikatora razlike između analiziranih grupa.
- 2) *Long-rank test* je ne-parametarski test bez pretpostavki o raspodeli preživljavanja. Omogućava poređenje krivih preživljavanja između grupa tako što poredi zabeleženi broj ishoda u svakoj grupi sa brojem ishoda koji bi se dobio u slučaju da je prihvaćena nulta hipoteza identičnih krivih preživljavanja (115).
- 3) *Cox-ov regresioni model proporcionalnog rizika* dozvoljava istovremenu analizu efekta više numeričkih i/ili kategoričkih faktora rizika na vreme preživljavanja. Omogućava kvantifikaciju efekta mereći razliku između grupa kao i prilagođavanje usled efekata nekontrolisanih varijabli (116).

Cox-ov model (117) proporcionalnog rizika je semi-parametarski regresioni model korišćen u medicinskim ispitivanjima preživljavanja za analizu rizika u toku vremena i za procenu efekta kovarijanti, tj. parametara za koje se smatra da utiču na vremena preživljavanja pacijenata. Osnovna pretpostavka Cox-ovog PH (engl. *Proportional Hazard*) modela je proporcionalnost rizika što rezultuje konstantnom stopom rizika HR koji se koristi za kvantifikaciju efekta kovarijanti a da pri tome nije neophodno poznavanje funkcije osnovnog rizika.

Cox-ov model preživljavanja je predstavljen vremenski zavisnom funkcijom rizika $h(t)$ koja zavisi od kovarijanti (prediktorskih varijabli, prediktora, prognostičkih faktora) x_i ($i = 1, 2, \dots, n$) na sledeći način

$$h(t) = h_o(t) \cdot e^{b_1x_1 + b_2x_2 + \dots + b_nx_n} \quad (4.6.5)$$

pri čemu t predstavlja vreme preživljavanja. Regresioni koeficijenti b_1, \dots, b_n mere veličinu efekta pojedinačnih prediktora i mogu se proceniti na osnovu eksperimentalnih podataka upotrebom parcijalne funkcije verodostojnosti. Sa $h_o(t)$ je opisana funkcija osnovnog rizika (engl. *baseline hazard*) koja odgovara funkciji rizika u trivijalnom slučaju kada su vrednosti svih prediktora jednake nuli. Glavni cilj primene Cox-ovog PH modela preživljavanja je identifikacija značajnih prognostičkih faktora, tj. detekcija kovarijanti koje značajno utiču na vreme preživljavanja i na procenu intenziteta njihovog efekta.

Najvažnija pretpostavka Cox-ovog PH modela je proporcionalnost rizika u posmatranim grupama. Rizik u toku vremena se menja ali na takav način da je odnos rizika u grupama konstantan. Odnos stopa rizika dvaju grupa HR je definisan sa

$$HR = \frac{h'(t)}{h(t)} \quad (4.6.6)$$

pri čemu funkcije $h'(t)$ i $h(t)$ uglavnom odgovaraju grupi većeg i manjeg rizika, respektivno. Uzimajući u obzir formulu (4.6.6) može se videti da je HR nezavistan od vremena

$$HR = \frac{h'(t)}{h(t)} = \frac{h_o(t) \cdot e^{\sum_{i=1}^n b_i x'_i}}{h_o(t) \cdot e^{\sum_{i=1}^n b_i x_i}} = e^{\sum_{i=1}^n b_i (x'_i - x_i)} = const.$$

Pretpostavimo li da postoji samo jedna dihotomna prediktorska varijabla koja može uzimati vrednosti 1=primenjen eksperimentalni lek i 0=placebo, proizilazi da je

$$HR = e^{b_1(1-0)} = e^{b_1}.$$

Ukoliko je $b_1 = 0$ onda je $HR = 1$ i zaključuje se da uzimanje leka nema efekta. Ukoliko je $b_1 > 0$ ($b_1 < 0$) onda je $HR > 1$ ($HR < 1$) i zaključuje se da je rizik za preživljavanje veći (manji) u eksperimentalnoj grupi u odnosu na kontrolnu.

Sumirajući prethodno, HR se može interpretirati na sledeći način:

- 1) $HR < 1$ znači da dolazi do *smanjenja rizika* pa se prediktori interpretiraju kao *dobri prognostički faktori*,
- 2) $HR = 0$ znači da nema razlike u posmatranim grupama,
- 3) $HR > 1$ znači da se prediktori interpretiraju kao *loši prognostički faktori* usled *povećanja rizika*.

U praktičnim primena Cox-ovog PH modela je uobičajeno da se pojavljuje relativno veliki broj kovarijanti povezanih sa ishodom. U selekciji optimalnog podskupa kovarijanti sa statistički najznačajnijim efektom na vreme preživljavanja u odnosu na sve moguće podskupove se primenjuju različiti algoritmi. Najčešće primenjivani su postupak izbora *unapred*, *unazad* i *postupak postupne selekcije*. U postupku izbora unapred (engl.*forward selection*) modelu se dodaje jedna po jedna kovarijanta, procenjuje se njen statistički značaj i na osnovu unapred prihvaćenog praga značajnosti donosi se odluka o zadržavanju ili odbacivanju. Postupak izbora unazad (engl.*backward selection*) je obrnut: početni model sadrži sve kovarijante koje se jedna po jedna eliminišu iz modela. Postupak postupne selekcije (engl.*stepwise selection*) je kombinacija prethodna dva. Nakon odabira najznačajnijih prediktorskih faktora i određivanja osnovne funkcije rizika i/ili preživljavanja moguće je proceniti verovatnoću da pacijent preživi duže od datog vremenskog perioda.

Za opravdanu primenu Cox-ovog modela je neophodno da bude zadovoljena pretpostavka o proporcionalnosti rizika. Provera proporcionalnosti se može izvršiti vizuelnom inspekcijom krivih preživljavanja zasnovanih na Kaplan-Meier modelu.

Upoređuju se log-log krive preživljavanja različitih grupa, ili se upoređuju očekivane i eksperimentalne krive preživljavanja. U slučaju proporcionalnosti ove krive su aproksimativno paralelne i nikako nije dozvoljeno da se seku. Objektivnija procena proporcionalnosti rizika je zasnovana na grafičkoj i/ili statističkoj analizi reziduala. Uglavnom se koriste: Schoenfeld reziduali za proveru proporcionalnosti, martingalni reziduali za proveru nelinearnosti, tj. prilagođenosti kovarijanti i reziduali odstupanja za proveru uticaja netipičnih rezultat (engl. *outliers*) (118).

4.7. Obrada rezultata

Procena prognostičke efikasnosti i analiza nezavisnosti računarskih, demografskih i kliničko-patoloških parametar izvršena je primenom ROC analize i Cox-ovog modela proporcionalnog rizika. Za krajnji ishod bolesti je definisana pojava metastaza kao direktna mera ishoda bolesti. Za analizu prognostičkih performansi, ROC analizu kao i za multivarijacionu Cox-ovu analizu korišćeni su programski paketi IBM SPSS Statistics v24 (IBM Corp. Armonk, NY) i Stata/MP13 (StataCorp, College Station, TX, USA).

AUC vrednosti su izračunate koristeći kontinuirane vrednosti dok su za Cox-ovu regresionu analizu korišćeni kategorizovani podaci. Za izračunavanje optimalnog praga kategorizacije je korišćen program X-tile 3.6.1 (Yale University, New Haven, USA) (119). Za proveru pretpostavke proporcionalnosti rizika korišćen je metod Schoenfeld-ovih reziduala (Stata/MP).

Interna validacija ROC i Cox-ove regresione analize za interval poverenja od 95% je izvršena primenom "bootstrap" interne validacije korišćenjem 1 000 preuzorkovanih podataka (120). Korekcija je izvršena korišćenjem programskih paketa Stata/MP i IBM SPSS Statistics.

Za tabelarni prikaz podataka i izradu grafika je korišćen program Excel, Microsoft Office Professional Plus 2016 kompanije Microsoft.

5. REZULTATI

Za procenu prognostičkog učinka starosti pacijentkinja i kliničko-patoloških parametara, kao i za procenu učinka teksturalnih i morfometrijskih parametara ispitivana je njihova moć da razvrstaju pacijente u odnosu na pojavu udaljenih metastaza što je ključni klinički događaj za ishod raka dojke. Kao kvantitativna mera prognostičkog učinka korišćena je AUC vrednost pri čemu je učinak rangiran shodno skali definisanoj u poglavlju 4.6.1. Efekat podele histograma i optimizacije praga binarizacije na prognostički učinak je kvantifikovan korišćenjem povećanja ili smanjenja u postignutim AUC vrednostima.

5.1. Kliničko-patološke karakteristike uzorka pacijentkinja

Analizirana grupa od 102 pacijentkinje je imala tumore manje veličine klasifikovane kao pT1 ili pT2⁷. Prema stepenu histološke diferencijacije tumora pacijenti su svrstani u gradus 1 i 2 sa samo jednom pacijentkinjom gradiranom kao G3. Ocena limfonodalnog statusa je ukazala da su svi pacijenti bez zahvaćenih limfnih čvorova i metastaza. Kod svih pacijentkinja iz uzorka izvršena je i analiza ekspresije estrogenskih i progesteronskih receptora (121). Analiza ER i PR receptora je ukazala na sledeću raspodelu rezultata unutar posmatrane grupe: 70 pacijentkinja (68.6%) su imale pozitivan nalaz na ER (srednja vrednost od 32 fmol/mL) dok je 32 pacijentkinja (38.2%) bilo pozitivno na PR (srednja vrednost 6 fmol/mL).

U posmatranoj grupi u vremenskom rasponu između 16 i 155 meseci (srednje vreme od 61 mesec) došlo je do razvoja metastaza kod 21 (20,6%) pacijentkinje. Metastaze su lokalizovane na jetri, plućima, kostima, koži i mišićima. Srednje vreme praćenja inverznom Kaplan-Meier-ovom metodom preostale 81 (80,4%) pacijentkinje je iznosilo 147 meseci sa ukupnim vremenskim rasponom od 77 do 165 meseci. Detaljne histopatološke i imunohistohemijske karakteristike posmatrane grupe pacijentkinja su prikazane u tabeli 5.1.1. Prikazane karakteristike su

⁷ pT1 odgovara najvećoj dimenziji tumora ≤ 2 cm, dok pT2 odgovara tumorima čija se najveća dimenzija kreće od 2 do 5 cm.

izračunate analizom rezultata celokupne grupe od 102 pacijentkinje i razvrstane su po kategorijama.

Tabela 5.1.1. Histopatološke i imunohistochemijske karakteristike analiziranog uzorka pacijenata

Parametar	Broj pacijenata	Metastaze (%) ^a
HER2 status		
HER2-	80	21
HER2+	22	20
ER status (prag=20 fmol/mg)		
ER-	32	13
ER+	70	25
PR status (prag=10 fmol/mg)		
PR-	63	21
PR+	39	20
Veličina tumora (cm)		
≤ 2	73	12
2 – 5	27	41
≥ 5	2	50
Limfno-nodalni status		
N0	102	21
N+	0	0
Histološki gradus		
G1	9	33
G2	92	17
G3	1	0
Broj raspršenih individualnih epitelnih ćelija^b		
Nizak (0 – 7,6)	60	5
Visok (8,3 – 37,0)	42	40

^aProcentualno iskazan broj metastaza u okviru podgrupe. ^bBroj raspršenih individualnih epitelnih ćelija za svakog pacijenta je procenjen kao nizak ili visok prema usvojenoj optimalnoj graničnoj vrednosti. Skraćenice: HER2 - humani epidermalni faktor rasta 2; ER - estrogenski receptori; PR - progesteronski receptori;

Tabela 5.1.2. Lokalizacija metastaza u analiziranoj grupi pacijentkinja

Mesto lokalizacije	Broja pacijentkinja	(%) ^a
Pluća	8	38
Kosti	7	33
Jetra	4	19
Koža	1	5
Mišići	1	5

^aProcentualno iskazan broj metastaza u odnosu na ukupan broj metastaza u analiziranoj grupi.

5.2. Prognošički učinak starosti pacijentkinja i kliničko-patoloških parametara

Parametri *Starost*, *Veličina tumora*, *ER* i *PR*, su u procesu statističke obrade podataka analizirani kao kontinuirane veličine da bi se izbeglo biranje graničnih vrednosti koje uvodi pristrasnost u procenu rezultata. Preostali parametri su prema svojoj prirodi kategoričke veličine. HER2+ odgovara pozitivnoj amplifikaciji HER2, HER2-obogaćen označava HER2+/PR-/ER- molekularni pod tip, dok se *Trostruko negativni* koristi kao odrednica za ER-/PR-/HER2- molekularnu pod kategoriju karcinoma dojke (122).

Tabela 5.2.1. Prognošički učinak starosti pacijenata i kliničko-patoloških parametara procenjen upotrebom AUC vrednosti

Parametar	AUC ^a	p-vrednost	95% IP
Starost (<i>godina</i>)	0,60	0,15	0,48 - 0,73
Veličina tumora (cm)	0,65 ^b	0,04*	0,51 - 0,78
Gradus	0,45	0,50	0,31 - 0,60
ER (<i>fmol/mg</i>)	0,60	0,14	0,46 - 0,75
PR (<i>fmol/mg</i>)	0,48	0,80	0,33 - 0,63
HER2+	0,49	0,45	0,32 - 0,67
HER2-obogaćen	0,47	0,67	0,33 - 0,61
Trostruko negativni	0,46	0,55	0,32 - 0,60

^aIzvršena "bootstrap" interna validacija; ^bAUC = 0,65 – maksimalna postignuta AUC vrednost kliničko-patoloških parametara; * $p \leq 0,05$. Skraćenice: AUC - površina ispod ROC krive; IP - interval poverenja; HER2 - humani epidermalni faktor rasta 2; ER - estrogenski receptori; PR - progesteronski receptori; HER2-obogaćen - ER-, PR-, HER2+ molekularni pod tip; Trostruko negativni - ER-PR- i HER2- molekularni pod tip.

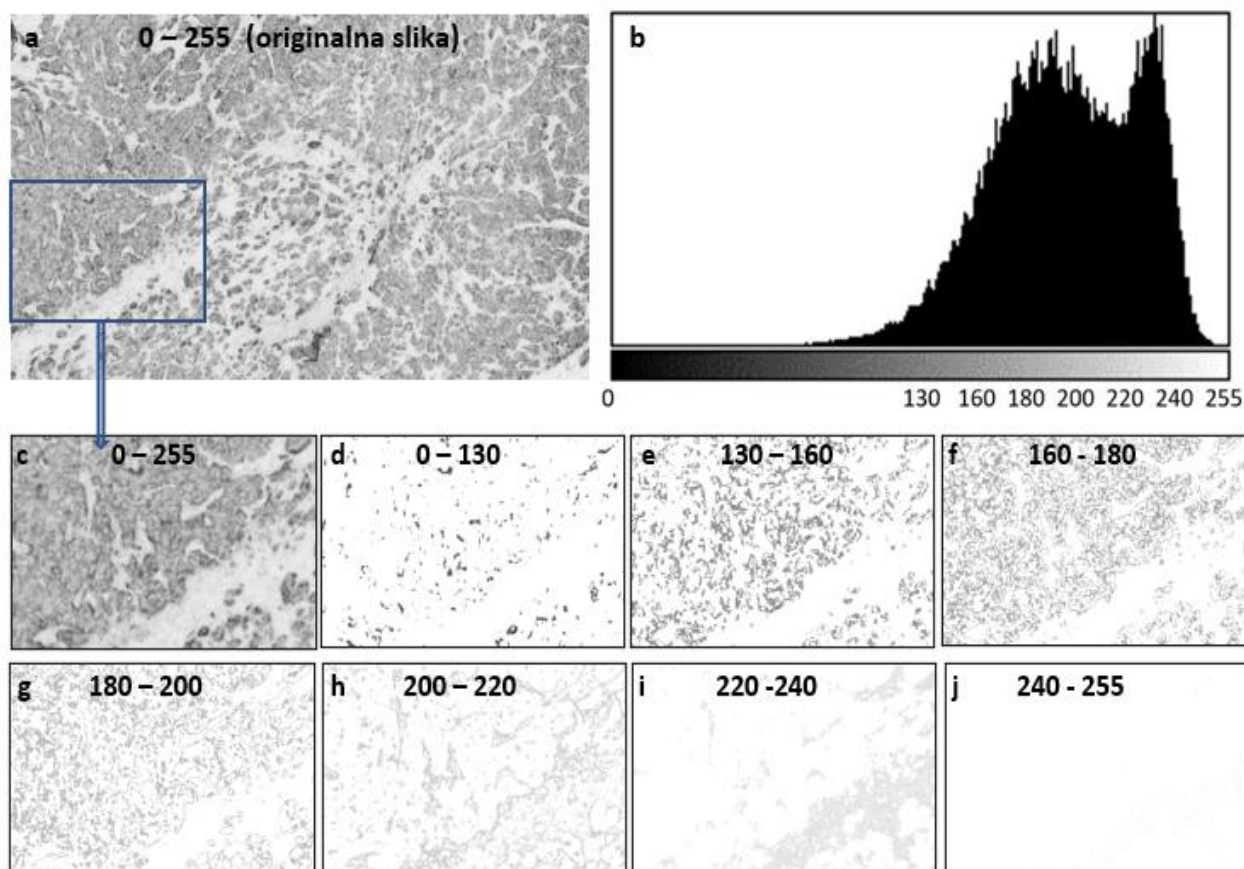
Za ispitivanu grupu pacijenata AUC vrednosti *Starosti* i kliničko-patoloških parametara su se kretale između 0,45 i 0,65. (tabela 5.2.1). Maksimalna postignuta AUC vrednost postignuta za kliničko-patološke prognošičke parametre je iznosila 0,65 i postignuta je za parametar *Veličina tumora*. Shodno skali definisanoj u poglavlju 4.6.1. ova vrednost je rangirana kao *dobra* a uz to je i statistički značajna sa $p=0,04$. Najlošiju vrednost je postigao *HER2+ status*.

Osim *Veličine tumora* od osam analiziranih kliničko-patoloških parametara samo su još dva, *Starost* (AUC = 0,60) i *ER* (AUC = 0,60), rangirani kao *dobri* ali nijedan nije postigao statističku značajnost ($p = 0,15$ i $p = 0,14$, respektivno). Preostalih pet kliničko-patoloških parametara: *Gradus*, *PR*, *HER2+*, *HER2-obogaćen* i *Trostruko negativni* takođe su bez statističkog značaja sa postignutim AUC vrednostima vrlo bliskim dijagonali (AUC = 0,5).

5.3. Raspodela intenziteta pan-citokeratinskog imunobojenja

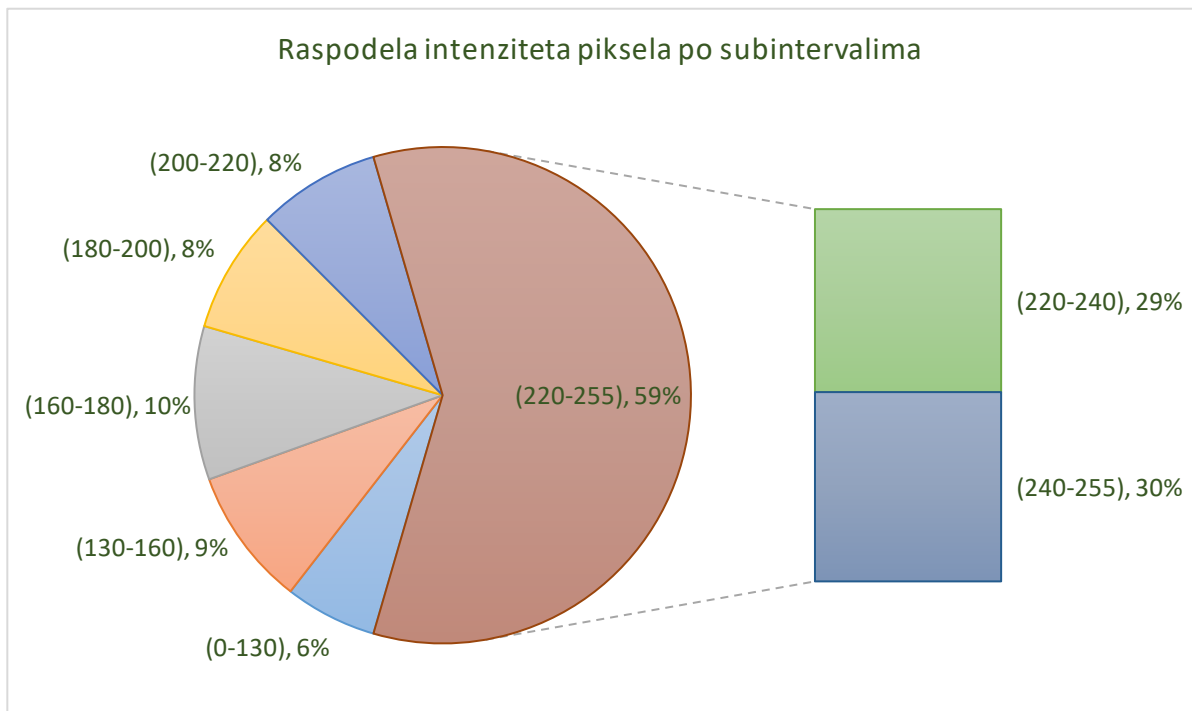
Analiza histograma je ukazala na neravnomernu raspodelu intenziteta imunobojenja u imunohistohemijskim preparatima. Stoga su pikseli na osnovu njihovog intenziteta podeljeni u nepreklapajuće podintervale. Vizuelni pregled histograma je ukazao da manjem broju piksela odgovaraju tamniji intenziteti niži od 150 pri čemu se većina piksela nalazi u svetlijem intervalu 150 - 250 . Da bi se ispitalo koji intenziteti pan-citokeratinskog bojenja imaju najveći prognostički značaj izvršena je podela celokupnog intervala intenziteta na podintervale: 0-130, 130-160, 160-180, 180-200, 200-220, 220-240 i 240-255 (slika 5.3.1.).

Podela originalnog intervala na sedam nepreklapajućih podintervala je rezultovalo sledećom prosečnom raspodelom intenziteta: 6% u prvom i najširem podintervalu najtamnijih piksela, 9% u podintervalu 130-160, 10% u podintervalu 160-180, 8% u 180-200, 8% u 200-220, 29% u 220-240 i 30% u poslednjem podintervalu 240-255 najsvetlijih piksela (slika 5.3.2).



Slika 5.3.1. Podela intervala intenziteta piksela 8-bitnih slika. (a) Originalna 8-bitna slika sa (b) pripadajućim histogramom: x-osa - intenzitet piksela, y-osa - broj piksela. Uveličanom segmentu originalne slike odgovara (c) celokupni interval intenziteta 0-255, (d) 0-130 podinterval, (e) 130 - 160, (f) 160 - 180, (g) 180 - 200, (h) 200 -220, (i) 220 - 240, i (j) 240 - 255 podinterval. Poslednja slika je skoro u potpunosti bela što je posledica činjenice da sadrži samo 1 307 od 122 516, tj. 1.07% ukupnog broja piksela originalne slike (c). Slike (e) i (f) sadrže skoro polovinu, tj. 23.6% i 22.8% od ukupnog broja piksela, respektivno.

Celokupni interval intenziteta 0 - 255 bi se mogao približno podeliti na tri podintervala intenziteta: visoki, srednji i niski. Podintervalu niskih intenziteta 0 - 130 pripadaju najtamniji pikseli kojima odgovaraju epitelne plaže najintenzivnije obojene pan-citokeratinom. U podintervalu 130 - 200 su raspoređeni pikseli srednjih intenziteta dok se u trećem podintervalu 200 - 255 nalaze najsvetliji pikseli kojima odgovaraju i najniži intenziteti pan-citokeratinskog imunobojenja.

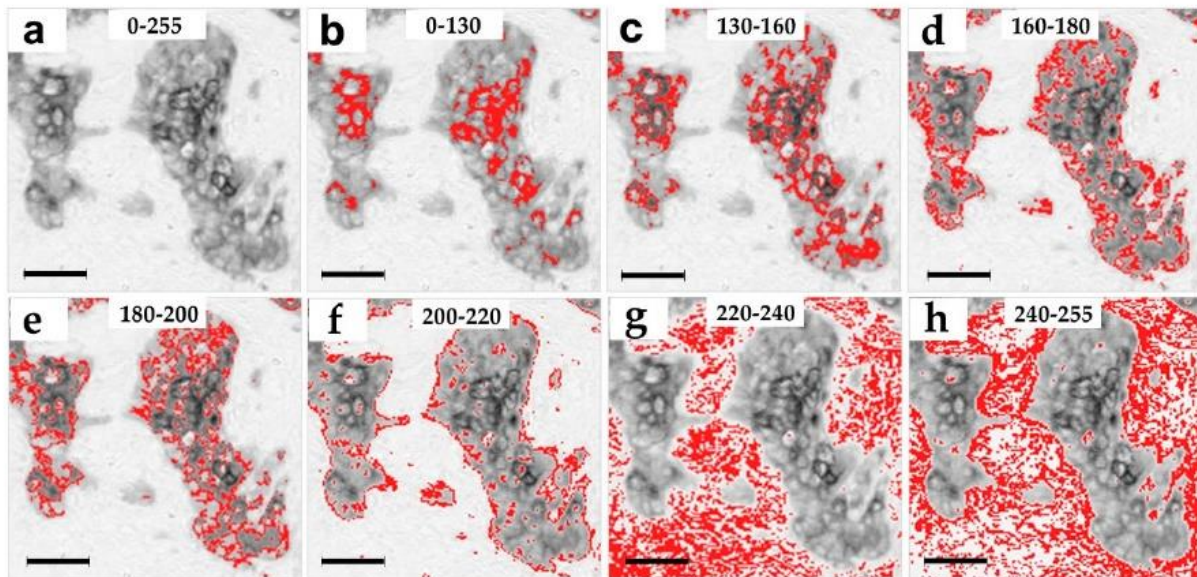


Slika 5.3.2. Raspodela piksela po podintervalima celokupnog intervala sivih tonova izražena u procentima. Više od polovine (59% piksela) su svetli pikseli intenziteta većih od 220.

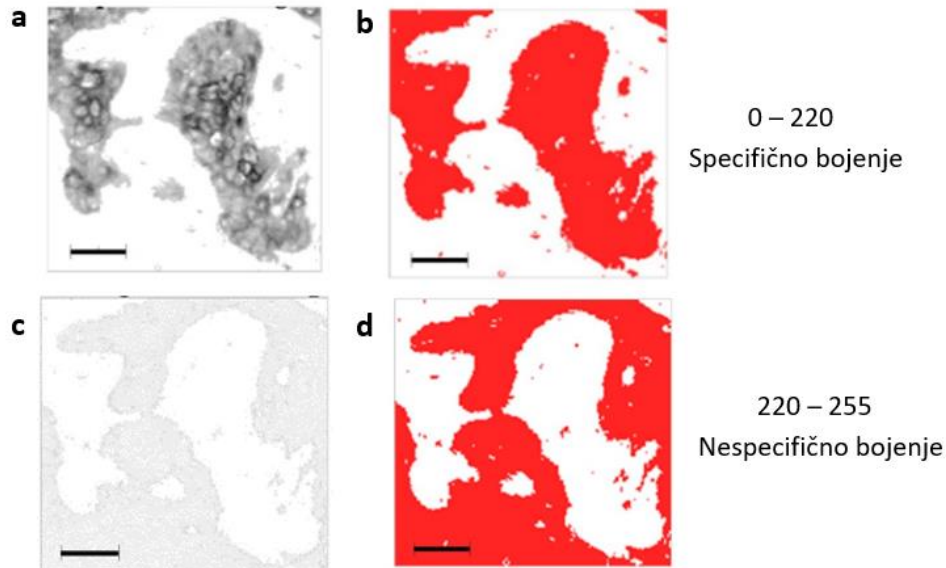
5.3.1. Specifično i nespecifično pan-citokeratinsko imunobojenje

Prostornu raspodelu intenziteta pan-citokeratinskog imunobojenja je lakše analizirati na slikama dobijenim preklapanjem originalne slike (slika 5.3.3. a) sa binarnom maskom na kojoj je pikselima iz pripadajućeg intervala intenziteta pridružena crvena boja (slika 5.3.3. b-h). Na tako dobijenim slikam se jasno vidi da su pikseli iz podintervala 0 - 220 raspoređeni unutar oblasti koje odgovaraju malignim plažama što je inače kompatibilno i sa raspodelom epitelnih ćelija.

Na osnovu uvida u detaljnu raspodele intenziteta bojenja u kombinaciji sa morfološko-patološkom analizom identifikovan je intenzitet 220 kao granica između specifičnog i nespecifičnog bojenja. Kompletan obrazac distribucije specifičnog bojenja unutar opsega 0 - 220 i nespecifičnog bojenja unutar opsega 220 - 255 kao i odgovarajuće binarne maske se može videti na slici 5.3.4.



Slika 5.3.3. Prostorna raspodela intenziteta pan-citokeratinskog imunobojenja. (a) Uvećani segment ($\times 320$) reprezentativne 8-bitne slike intervala intenziteta 0 - 255, (b) originalna slika sa pikselima iz intervala 0 - 130 obojenim u crveno čime je omogućena vizualizacija raspodele visokih intenziteta bojenja, srednji intenziteti (c) 130 - 160, (d) 160 - 180, (e) 180 - 200 i intervali niskog intenziteta bojenja (f) 200 - 220, (g) 220 - 240 i (h) 240 - 255. Veličina piksela: 1.8 μm . Skala: 50 μm .



Slika 5.3.4. Specifično i nespecifično bojenje pan-citokeratinom. (a) Uvećani segment ($\times 320$) reprezentativne 8-bitne slike koji sadrži samo specifične piksele iz intervala intenziteta 0 - 220 i (b) pripadajuća binarna maska (c) Nespecifično obojeni segment slike iz podintervala 220 - 255 i (d) pripadajuća binarna maska. Veličina piksela: 1.8 μm . Skala: 50 μm .

5.4. Prognoštički učinak teksturalnih parametra

Analiza teksture je izvršena na 8-bitnim slikama sa intenzitetom sivih tonova koji se menjao u intervalu od 0 (crno) do 255 (belo) i na sedam setova slika dobijenih podelom celokupnog intervala intenziteta na podintervale, kao što je objašnjeno u poglavlju 4.4.1.

Podela histograma je rezultovala povećanjem broja belih piksela na slici usled konverzije svih piksela izvan zadatog podintervala u belu boju (slika 5.3.1. d-j). GLCM analiza koja je rađena na segmentiranim slikama uzela je u obzir piksele iz pripadajućeg podintervala kao i nepripadajuće piksele intenziteta 255 koji su u prethodnom koraku konvertovani u bele. Ovi pikseli nisu uzeti u obzir prilikom izračunavanja parametara statistike prvog reda zato što ova analiza može da isključi bele piksele.

U tabeli 5.4.1. su prikazani rezultati samo onih parametara koji su ostvarili *značajan prognoštički učinak* ($AUC \geq 0,60$) u bar jednom od podintervala uključujući i celokupni interval intenziteta 0 - 255. Zarad jasnijeg poređenja postignutih AUC vrednosti i procene efekta podele histograma vredi ponoviti da je polazna tačka u AUC prostoru dijagonala ($AUC = 0,50$) kojoj odgovara odsustvo diskriminacije i da se vrednosti ispod dijagonale mogu invertovati kao što je to objašnjeno u poglavlju 4.6.1. Tako su, na primer, parametri koji su postigli AUC vrednosti 0,66 i 0,34 ekvivalentni po učinku.

Najvišem ostvarenom prognoštičkom učinku teksturalnih parametra rangiranim kao *vrlo dobar* odgovara statistički značajna AUC vrednost od 0,77 postignuta za parametar ASM u intervalu 0 - 255. Ukupno 43 postignute AUC vrednosti su rangirane ili kao *dobre* ili kao *vrlo dobre*. Dvadeset i jedna AUC vrednost postignuta za teksturalne parametre je nadmašila maksimalnu vrednost od 0,65 postignutu za analizirane kliničko-patološke parametre (tabela 5.2). Šest postignutih AUC vrednosti su rangirane kao *vrlo dobre* sa $AUC \geq 0,70$ ili $AUC \leq 0,30$.

Parametar *Površina* je u pet podintervala postigao prognoštički značajne AUC vrednosti od čega tri nadmašuju maksimalnu AUC vrednost postignutu za kliničko-patološke parametre sa statistički značajnim ($p \leq 0,05$) AUC vrednostima od 0,69 i

0,31 postignutim u podintervalima 130 - 160, 160 - 180 i 220 - 240 (slika 5.4.1). Najlošiji prognostički učinak *Površine* koji se smatra za *nedovoljno dobar* je ostvaren u podintervalu 200 - 220. Uticaj podele histograma na ovaj parametar je pozitivan i iznosi $\Delta AUC = 0,17$ (tabela 5.4.2). Pri tome je $\Delta AUC = AUC_{max} - AUC_{min}$ računat tako da što su sve vrednosti ispod dijagonale invertovane da bi se omogućilo poređenje.

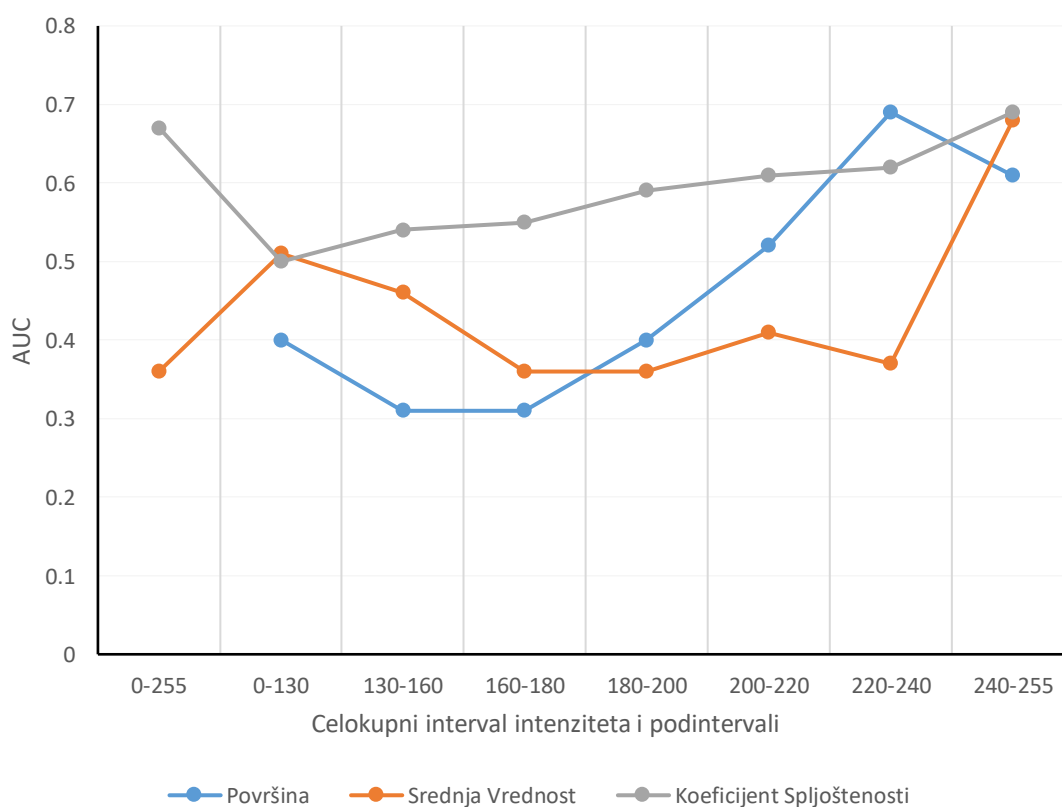
Tabela 5.4.1. AUC vrednosti teksturalnih parametara postignute na celokupnom intervalu intenziteta piksela i u podintervalima

		AUC ^c p-vrednost 95% IP							
		0-255	0-130	130-160	160-180	180-200	200-220	220-240	240-255
PARAMETRI STATISTIKE PRVOG REDA	Površina	-	0,40 ^b	0,31 ^{b,c}	0,31 ^{b,c}	0,40	0,52	0,69 ^{b,c}	0,61 ^b
		-	0,16	0,006*	0,007*	0,17	0,77	0,009*	0,13
		-	0,26-0,55	0,18-0,43	0,19-0,43	0,27-0,53	0,38-0,66	0,67-0,80	0,48-0,74
	Srednja vrednost	0,36 ^b	0,51	0,46	0,36 ^b	0,35 ^{b,c}	0,41	0,37 ^b	0,68 ^{b,c}
		0,05*	0,91	0,58	0,05*	0,03*	0,21	0,007*	0,01*
		0,23-0,49	0,27-0,66	0,32-0,62	0,23-0,50	0,22-0,47	0,27-0,54	0,24-0,50	0,56-0,80
	Koeficijent sploštenosti	0,67 ^{b,c}	0,50	0,54	0,55	0,59	0,61 ^b	0,62 ^b	0,69 ^{b,c}
		0,02*	0,98	0,52	0,52	0,20	0,12	0,10	0,01*
		0,55-0,78	0,35-0,64	0,31-0,61	0,38-0,69	0,45-0,72	0,48-0,74	0,49-0,75	0,57-0,81
PARAMETRI GLCM ANALIZE	IDM	0,75 ^{b,c}	0,61 ^b	0,66 ^{b,c}	0,68 ^{b,c}	0,64 ^b	0,53	0,41	0,48
		0,000*	0,14	0,03*	0,01*	0,06	0,70	0,20	0,75
		0,64-0,85	0,44-0,74	0,56-0,81	0,56-0,81	0,50-0,77	0,39-0,67	0,27-0,51	0,34-0,61
	ASM	0,77 ^{b,c}	0,61 ^b	0,66 ^{b,c}	0,68 ^{b,c}	0,64 ^b	0,52	0,36 ^b	0,45
		0,000*	0,14	0,03*	0,01*	0,06	0,70	0,04*	0,53
		0,66-0,87	0,44-0,70	0,56-0,81	0,56-0,81	0,50-0,77	0,38-0,67	0,22-0,46	0,38-0,59
	CON	0,39 ^b	0,38 ^b	0,35 ^{b,c}	0,31 ^{b,c}	0,34 ^{b,c}	0,42	0,45	0,54
		0,13	0,08	0,04*	0,01*	0,03*	0,28	0,47	0,55
		0,28-0,53	0,24-0,54	0,19-0,43	0,19-0,43	0,22-0,48	0,29-0,56	0,33-0,58	0,42-0,68
	ENT	0,28 ^{b,c}	0,40 ^b	0,34 ^{b,c}	0,32 ^{b,c}	0,37 ^b	0,48	0,65 ^{b,c}	0,54
		0,002*	0,17	0,03*	0,01*	0,06	0,81	0,04*	0,59
		0,19-0,40	0,27-0,57	0,20-0,45	0,20-0,45	0,34-0,50	0,34-0,63	0,56-0,79	0,41-0,67
COR	0,60 ^b	0,61 ^b	0,66 ^{b,c}	0,71 ^{b,c}	0,71 ^{b,c}	0,64 ^b	0,56	0,70 ^{b,c}	
	0,17	0,14	0,02*	0,004*	0,004*	0,06	0,43	0,007*	
	0,43-0,75	0,43-0,74	0,60-0,82	0,60-0,82	0,58-0,83	0,52-0,76	0,43-0,69	0,58-0,81	
^d AUC		0,68	0,61	0,66	0,68	0,66	0,62	0,65	0,67

^aIzvršena "bootstrap" interna validacija; ^bAUC ≥ 0,60 - prognostički značajne vrednost; ^cAUC ≥ 0,65 (AUC = 0,65 – maksimalnu AUC vrednost postignutu za kliničko-patološke parametre); ^dAUC - srednja vrednost prognostički značajnih postignutih AUC vrednosti; *p ≤ 0,05. Skraćenice: AUC - površina ispod ROC krive; IP - interval poverenja;

AUC vrednosti postignute za parametar *Srednja vrednost* intenziteta piksela su se kretale od najslabije AUC = 0,51 (podinterval 0 - 130) do maksimalne vrednosti AUC = 0,68 (podinterval 240 - 255) statistički značajne (slika 5.4.1). *Srednja vrednost* je ostvarila značajan prognostički učinak u celokupnom intervalu intenziteta i u četiri podintervala pri čemu je jedna postignuta AUC vrednost izjednačila a jedna nadmašila maksimalnu postignutu AUC vrednost za kliničko-patološke parametre. Ostvarena su povećanja $\Delta AUC = 0,17$ i $\Delta AUC^* = 0,04$. Povećanje ostvareno u odnosu na originalne 8-bitne slike je računato kao razlika maksimalne AUC vrednosti postignute u podintervalima i vrednosti ostvarene u intervalu 0-255, tj. $\Delta AUC^* = 0,68 - 0,64 = 0,04$. U slučaju parametra *Srednja vrednost*, AUC postignut na celokupnom intervalu intenziteta je 0,36 što je nakon inverzije dalo AUC = 0,64.

Uticaj podele histograma na AUC vrednosti statističkih parametara prvog reda



Slika 5.4.1. Prognostički učinak parametara statistike prvog reda u celokupnom intervalu intenziteta piksela i u podintervalima. Učinak je analiziran i kvantifikovan primenom ROC-AUC analize.

Postignute AUC vrednosti parametra *Koeficijent spljoštenosti* su se menjale od 0,50 (podinterval 0 - 130) do statistički i prognostički značajne vrednosti od 0,69 (podinterval 240 - 255). U celokupnom intervalu intenziteta i u tri podintervala *Koeficijent spljoštenosti* je postigao prognostički značajan učinak pri čemu maksimalna postignuta AUC vrednost i vrednost postignuta u intervalu 0 - 255 prevazilaze maksimalnu AUC vrednost postignutu za kliničko-patološke parametre. Ostvarena povećanja su $\Delta AUC = 0,19$ i $\Delta AUC^* = 0,02$.

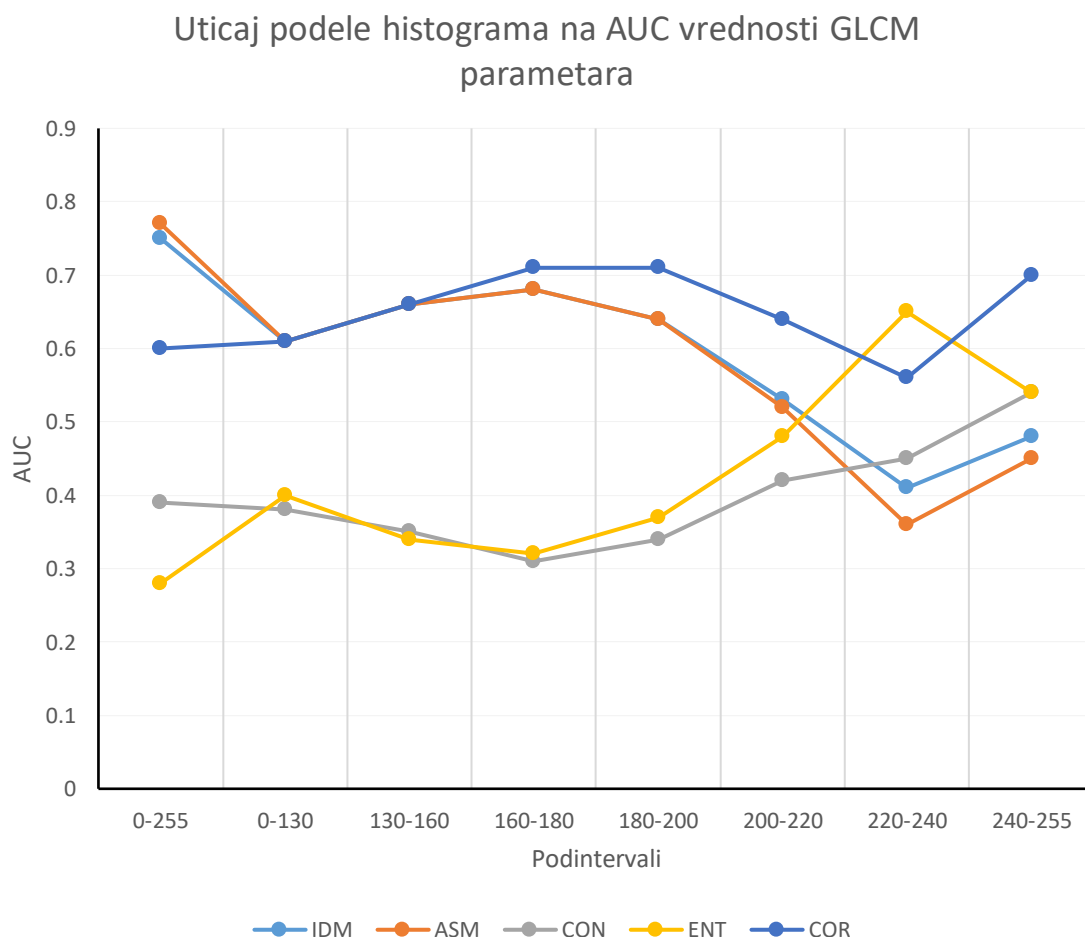
Parametar IDM je postigao prognostički i statistički značajnu AUC vrednost (AUC = 0,75) na celokupnom intervalu intenziteta 0 - 255. U četiri podintervala učinak IDM je rangiran kao *dobar* pri čemu su u podintervalima 130 - 160 i 160 - 180 postignute AUC vrednosti 0,66 i 0,68, respektivno (slika 5.4.2), koje su rangirane kao *dobre* i nadmašuju maksimalnu AUC vrednost postignutu za kliničko-patološke parametre. Podela histograma je imalo dvojak efekat na parametar IDM sa doprinosima od $\Delta AUC = 0,16$ i $\Delta AUC^* = -0,07$. Negativna vrednost ΔAUC^* znači da je došlo do redukcije u postignutim AUC vrednostima na podintervalima u odnosu na celokupni interval intenziteta.

Tri postignute AUC vrednosti parametra ASM nadmašuju maksimalnu AUC vrednost postignutu za kliničko-patološke parametre. Najviša AUC = 0,77 je postignuta u intervalu 0 - 255 dok su nešto niže vrednosti, AUC = 0,66 i 0,68, postignute u podintervalima 130 - 160 i 160 - 180, respektivno. Ukupno je postignuto šest prognostički značajnih vrednosti pri čemu je zabeleženo i $\Delta AUC = 0,16$ i $\Delta AUC^* = -0,09$.

Dve statistički značajne (na nivou značajnosti $p \leq 0,05$) AUC vrednosti postignute za parametar CON, AUC = 0,31 i AUC = 0,34, nadmašuju maksimalnu AUC vrednost postignutu za kliničko-patološke parametre. Ostvarene su u podintervalima 160 - 180 i 280 - 200, respektivno. Podela histograma je za CON rezultovalo i povećanjima u učinku od $\Delta AUC = 0,14$ i $\Delta AUC^* = 0,08$.

Od šest prognostički značajnih AUC vrednosti postignutih za parametar ENT tri su statistički značajne (na nivou značajnosti $p \leq 0,05$) i premašuju maksimalnu AUC vrednost postignutu za kliničko-patološke parametre: AUC = 0,28 (*vrlo dobar*, podinterval 0 - 255), 0,34 (*dobar*, podinterval 130 - 160) i 0,32 (*dobar*, podinterval 160 - 180). Ostvarena su i povećanja od AUC = 0,14 i $\Delta AUC^* = -0,04$.

Parametar COR je postigao sedam prognostički značajnih AUC vrednosti. U četiri podintervala postiže statistički značajne vrednosti čiji prognostički učinak premašuje maksimalnu AUC vrednost postignutu za kliničko-patološke parametre: AUC = 0,66 (*dobar*, podinterval 130 - 160), 0,71 (*vrlo dobar*, podintervali 160 - 180 i 180 - 200) i 0,70 (*vrlo dobar*, podinterval 240 - 255). Ostvarena su i povećanja $\Delta AUC = 0,15$ i $\Delta AUC^* = 0,11$.



Slika 5.4.2. Prognostički učinak teksturalnih parametara GLCM analize u analiza celokupnom intervalu intenziteta piksela i u podintervalima.. Učinak je analiziran i kvantifikovan primenom ROC-AUC analize.

Svi ispitivani parametri su postigli prognostički značajan učinak u celokupnom intervalu intenziteta i u bar tri podintervala (slike 5.4.1 i 5.4.2). Prema broju postignutih prognostički značajnih vrednosti ($AUC \geq 0,60$ ili $AUC \leq 0,40$) *najmanje uspešan parametar je Koeficijent spljoštenosti* sa značajnim vrednostima postignutim samo u tri od sedam podintervala. Slede ga *Površina, Srednja vrednosti, IDM i CON* sa

prognostički značajnim AUC vrednostima postignutim u četiri podintervala, i ASM i ENT sa značajnim AUC vrednostima postignutim u pet podintervala. Prema broju podintervala u kojima su postignute prognostički značajne AUC vrednosti, *najuspešniji parametar je COR* sa šest podintervala pri čemu je maksimalna AUC vrednost postignuta za kliničko-patološke parametre nadmašena četiri puta.

Od svih postignutih prognostički značajnih AUC vrednosti šest su rangirane kao *vrlo dobre* i sve su postignute za GLCM parametre: tri u celokupnom intervalu intenziteta uključujući i maksimalnu vrednost AUC = 0,77 postignutu za parametar ASM, dok su tri preostale *vrlo dobre* AUC vrednosti postignute za parametar COR (slika 5.4.2).

5.4.1. Prognostički učinak teksturalnih parametara u intervalima specifičnog i nespecifičnog pan-citokeratinskog imunobojenja

AUC vrednosti ostvarene u intervalima specifičnog i nespecifičnog pan-citokeratinskog imunobojenja za parametre statistike prvog reda i GLCM parametara su prikazane u tabeli 5.4.2.

Prema broju prognostički značajnih postignutih AUC vrednosti učinak razmatranih parametara u intervalima 0 - 220 i 220 - 255 je izjednačen. U intervalu specifičnog bojenja 0 - 220 četiri parametra su ostvarila prognostički značajan učinak. Postignute AUC vrednosti su statistički značajne (na nivou značajnosti $p \leq 0,05$), rangirane su kao *dobre* i sve četiri premašuju maksimalnu AUC vrednost postignutu za kliničko-patološke parametre. Najviša postignuta AUC vrednost⁸ u intervalu 0 - 220 je 0,68. Postignuta je tri puta i sva tri puta je ostvarena za GLCM parametre.

U intervalu nespecifičnog bojenja 220 - 255 su zabeleženi nešto slabiji rezultati. Iako su takođe četiri parametra postigla prognostički značajne AUC vrednosti samo AUC = 0,68 postignut za parametar *Površina* premašuje maksimalnu AUC vrednost postignutu za kliničko-patološke parametre.

⁸ Kao što je ranije objašnjeno, AUC = 0,32 dobijen za parametar ENT je ekvivalentan sa AUC = 0,68.

Tabela 5.4.2. AUC vrednosti teksturalnih parametara postignuti u intervalima specifičnog i nespecifičnog bojenja

AUC ^a p-vrednost 95% IP								
	Površina	Srednja Vrednost	Koeficijent Spljoštenosti	IDM	ASM	CON	ENT	COR
0-220	0,33 ^{b,c}	0,57	0,56	0,68 ^{b,c}	0,68 ^{b,c}	0,41	0,32 ^{b,c}	0,59
	0,02*	0,34	0,68	0,01*	0,01*	0,22	0,01*	0,21
	0,20-0,45	0,40-0,73	0,37-0,79	0,57-0,79	0,57-0,78	0,30-0,53	0,21-0,44	0,44-0,74
220-255	0,68 ^{b,c}	0,55	0,64 ^b	0,41	0,36 ^b	0,42	0,61 ^b	0,56
	0,01*	0,48	0,05*	0,21	0,04*	0,28	0,12	0,37
	0,57-0,79	0,42-0,69	0,51-0,77	0,29-0,53	0,24-0,47	0,30-0,55	0,49-0,73	0,43-0,71

^aizvršena "bootstrap" interna validacija; ^bAUC \geq 0,60 (prognostički značajni parametri), ^cAUC \geq 0,65; * $p \leq$ 0,05. Skraćenice: AUC - površina ispod ROC krive; IP - interval poverenja;

5.4.2. Poboljšanje prognostičkog učinka teksturalnih parametara usled podele histograma i procena prognostički najefikasnijeg intervala intenziteta imunobojenja

Svi razmatrani parametri su poboljšali prognostički učinak usled podele histograma uz prosečno povećanje u ostvarenim AUC vrednostima u podintervalima od $\Delta AUC = 0,148$ (tabela 5.4.3). Najveće povećanje $\Delta AUC = 0,69 - 0,50 = 0,19$ je zabeleženo za parametar *Koeficijent spljoštenosti*. Inače su se statistički parametri prvog reda u odnosu na GLCM parametre pokazali osetljiviji na podelu histograma sa prosečnim povećanjima od $\Delta AUC = 0,177$ i $\Delta AUC = 0,15$, respektivno.

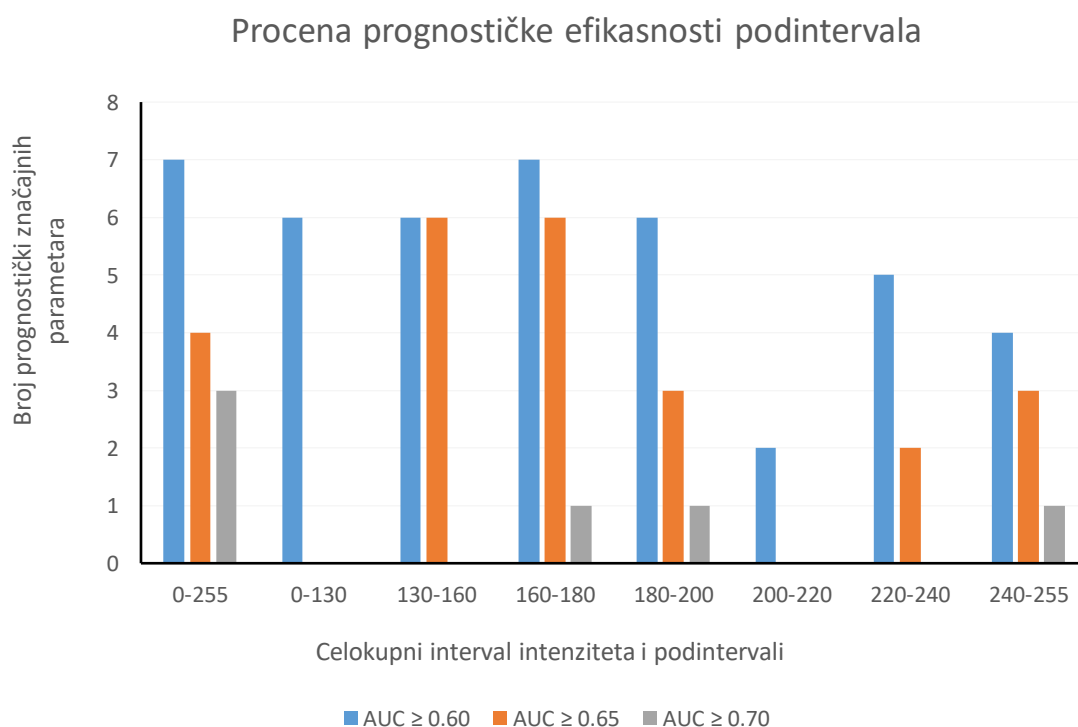
Tabela 5.4.3 Efekat podele histograma na prognostički učinak teksturalne parametre

	Površina	Srednja vrednost	Koeficijent spljoštenosti	IDM	ASM	CON	ENT	COR
ΔAUC	0,17	0,17	0,19	0,16	0,16	0,14	0,14	0,15
ΔAUC^*	n/a	0,04	0,02	-0,07	-0,09	0,08	-0,04	0,11

$\Delta AUC = AUC_{max} - AUC_{min}$ – doprinos postignut u podintervalima; ΔAUC^* - maksimalan doprinos postignut u podintervalima u odnosu na celokupni interval intenziteta. Skraćenice: AUC - površina ispod ROC krive;

U slučaju četiri parametra podela histograma je dovelo i do poboljšanja prognostičkog učinka u odnosu na učinak u celokupnom intervalu intenziteta.

Prema broju parametar koji su ostvarili prognostički značajan učinak najefikasniji su celokupni interval intenziteta i podinterval 160 - 180 u kojima čak sedam parametara postiže $AUC \geq 0,60$ (slika 5.4.3). Međutim, dok u celokupnom intervalu intenziteta samo četiri parametra postižu AUC vrednosti koje nadmašuju maksimalnu AUC vrednost postignutu za kliničko-patološke parametre, na *prognostički najefikasnijem intervalu* 160 - 180 čak šest parametara ostvaruje $AUC > 0,65$ sa srednjom postignutom AUC vrednošću od 0,681 (tabela 5.4.1). Slična efikasnost je ostvarena i u intervalu 130 - 160 na kome šest parametara postižu AUC vrednosti veće od maksimalne AUC vrednosti postignute za kliničko-patološke parametre. Najskromniji učinak je postignut u podintervalu 200 - 220. Samo dva parametra su ostvarila prognostički značajne AUC vrednosti. Rangirani su kao *dobri* ali nisu uspeali da dostignu maksimalni učinak standardno korišćenih parametara.



Slika 5.4.3. Procena prognostički najefikasnijeg intervala - procena je izvršena ustanovljavanjem i poređenjem broja parametara prognostički značajnog učinka ($AUC \geq 0,60$) u celokupnom intervalu intenziteta i u podintervalima.

5.5. Prognostički učinak morfometrijskih parametara epitelnih malignih plaža

Varijacija praga binarizacije se potvrdila kao kritičan korak za poboljšanje prognostičkog učinka morfometrijskih parametara usled efekta koje je imala na AUC vrednosti. Na primer, AUC vrednosti postignute za parametar oblika *Cirkularnost* su se u funkciji praga binarizacije menjale od 0,50 za Prag 170 do 0,26 za Prag 240 (tabela 5.5.1). Na slici 5.5.1 je dat primer rezultata dobijenog procesom automatske binarizacije (slika b) dok su preostale slike rezultat korisnički podešenog praga na različite vrednosti. Primenom najnižeg korišćenog praga dobijena je izrazito svetla slika što je i bilo za očekivati s obzirom da histogram originalne slike sadrži vrlo mali broj piksela ispod ove vrednosti intenziteta. Daljom varijacijom praga su generisane tamnije slike. Slike dobijene korišćenjem pragova 220 i 240 su vrlo nalik slici dobijenoj u procesu automatske binarizacije što je i razumljivo uzimajući u obzir da je za sliku odabranu za primer u automatskom procesu korišćen prag 226 (slika 5.5.1 c).

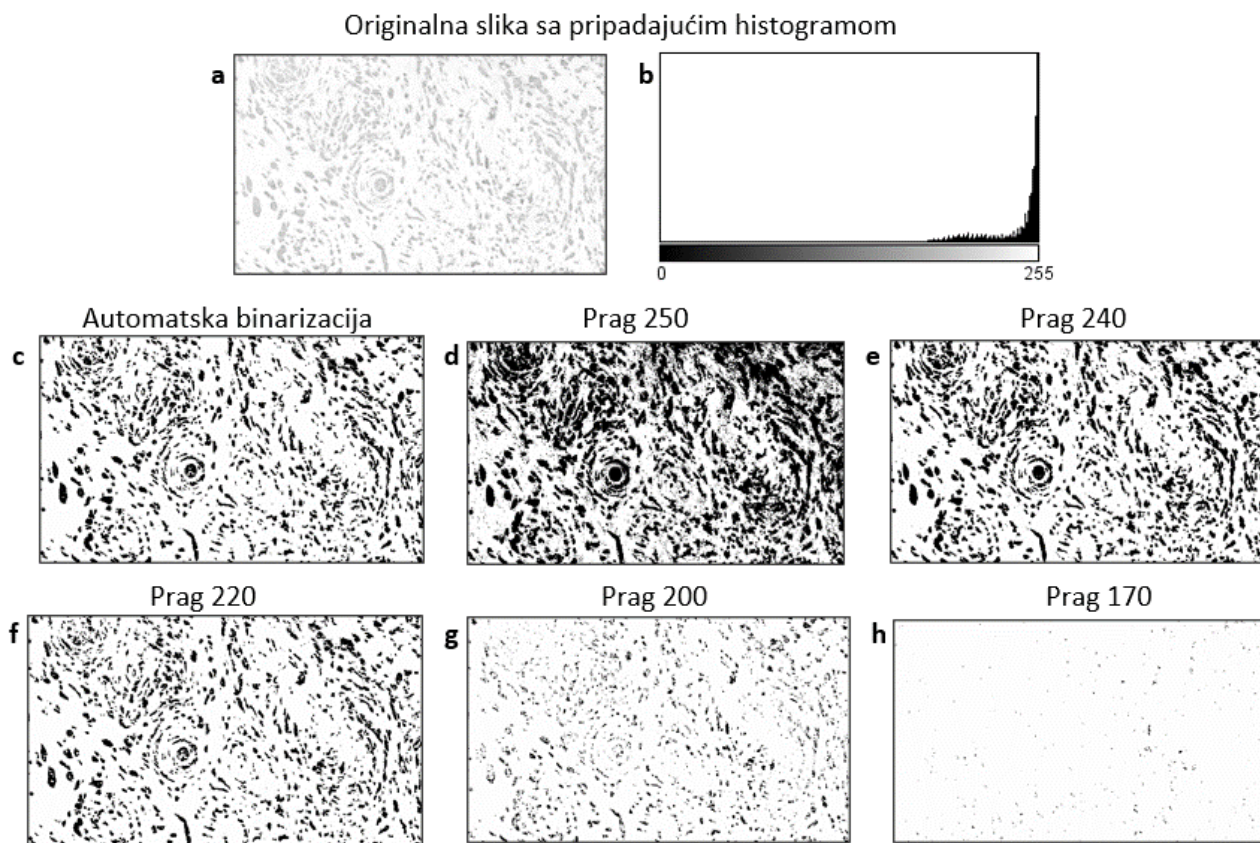
5.5.1. Prognostički učinak morfometrijskih parametara u funkciji praga binarizacije

Optimizacijom praga binarizacije je dobijeno šest setova slika, jedan procesom automatske binarizacije a preostalih pet korisnički odabirom praga. Za svaki pojedinačni set slika je izvršena procena prognostičkog učinka izračunatih morfometrijskih parametara.

U cilju ispitivanja efekta optimizacije procesa binarizacije, vrednosti praga su menjane bez primene filtera za morfološke karakteristika epitelnih klastera. Filter za cirkularnost je bio podešena u interval 0,0 - 1,0 a filter za veličinu na 0,0 - beskonačno što je praktično značilo da su uzete u obzir sve epitelne plaže.

U odsustvu filtriranja *cirkularnosti* i *veličine* postignute AUC vrednosti su se menjale između 0,50 i 0,25. Maksimalni prognostički učinak je ostvaren na pragu binarizacije 240 za parametar *Solidnost* sa statistički značajnom (na nivou značajnosti $p \leq 0,05$) postignutom AUC vrednošću od 0,25 što je ekvivalentno sa

AUC = 0,75, kao što je objašnjeno u poglavlju 4.6.1. Najlošije postignute AUC vrednosti bez prognostičkog značaja su ostvarene za parametre *Broj čestica* i *Cirkularnost* za pragove 250 i 170, respektivno.



Slika 5.5.1 Efekat optimizacije praga binarizacije. Binarne slike dobijena od (a) originalne 8-bitne slike procesom (c) automatske binarizacije (Prag 226) i korisničkim podešavanjem praga na (d) Prag 250, (e) Prag 240, (f) Prag 220, (g) Prag 200 i (h) Prag 170.

Ukupno 21 postignuta AUC vrednost je bila prognostički značajna pri čemu je 12 postignutih vrednosti nadmašilo maksimalnu AUC vrednost postignutu za kliničko-patološke parametre (AUC = 0,65 postignut za *Veličina tumora*, tabela 5.2.1). Devet postignutih AUC vrednosti su zadovoljile $AUC \geq 0,70$ ili $AUC \leq 0,30$ i rangirane su kao *vrlo dobre*.

Tabela 5.5.1 AUC vrednosti morfometrijskih parametrov epitelnih malignih plaž postignute pri avtomatski binarizaciji i korisnički odabranim pragovima binarizacije^a

AUC ^b						
95% IP						
p-vrednost						
Parametar	AutoBin	Prag 250	Prag 240	Prag 220	Prag 200	Prag 170
Broj čestica	0,46	0,50	0,41	0,51	0,31 ^{c,d}	0,38 ^c
	0,33-0,59	0,37-0,63	0,28-0,53	0,37-0,66	0,19-0,42	0,25-0,51
	0,59	0,98	0,18	0,85	0,01*	0,08
Ukupna površina	0,45	0,63 ^c	0,57	0,40 ^c	0,37 ^c	0,43
	0,32-0,59	0,51-0,76	0,44-0,71	0,26-0,55	0,23-0,52	0,28-0,58
	0,52	0,06	0,31	0,18	0,07	0,35
Prosečna veličina	0,52	0,74 ^{c,d}	0,67 ^d	0,45	0,47	0,51
	0,38-0,66	0,64-0,84	0,54-0,79	0,30-0,60	0,31-0,62	0,37-0,65
	0,80	0,00*	0,02*	0,48	0,62	0,88
Cirkularnost	0,33 ^{c,d}	0,63 ^c	0,26 ^{c,d}	0,36 ^c	0,54	0,50
	0,20-0,45	0,51-0,75	0,15-0,36	0,24-0,48	0,38-0,69	0,36-0,65
	0,01*	0,06	0,00*	0,05*	0,62	0,96
Solidnost	0,28 ^{c,d}	0,61 ^c	0,25 ^{c,d}	0,28 ^{c,d}	0,51	0,47
	0,17-0,40	0,48-0,74	0,15-0,34	0,18-0,39	0,36-0,65	0,33-0,62
	0,00*	0,12	0,00*	0,00*	0,94	0,69
Feret dijametar	0,60 ^c	0,71 ^{c,d}	0,74 ^{c,d}	0,52	0,48	0,52
	0,46-0,71	0,61-0,81	0,63-0,84	0,37-0,68	0,32-0,63	0,37-0,66
	0,17	0,00*	0,00*	0,75	0,77	0,83
MinFeret	0,61 ^c	0,73 ^{c,d}	0,74 ^{c,d}	0,52	0,49	0,53
	0,48-0,74	0,64-0,83	0,64-0,85	0,37-0,68	0,33-0,64	0,39-0,67
	0,12	0,00*	0,00*	0,73	0,86	0,67
^e \overline{AUC}	0,68	0,68	0,73	0,65	0,66	0,52

^aNisu primenjeni filteri za *cirkularnost* i *veličinu* tako da su u obzir uzete sve epitelne plaže.

^bIzvršena "bootstrap" interna validacija; ^cAUC \geq 0,60 - prognostički značajne vrednost,

^dAUC \geq 0,65 (AUC = 0,65 - maksimalnu AUC vrednost postignutu za kliničko-patološke parametre); ^e \overline{AUC} - srednja vrednosti prognostički značajnih postignutih AUC vrednosti;

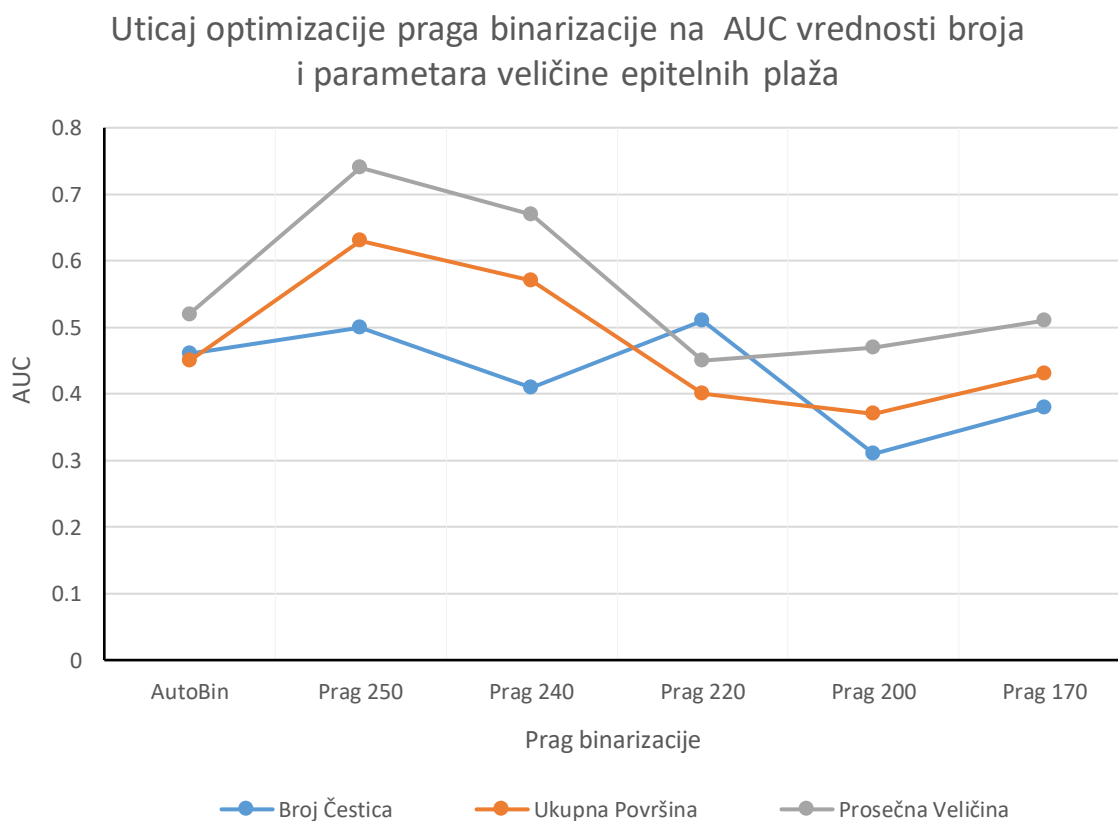
* $p \leq 0,05$. Skraćenice: AUC - površina ispod ROC krive; IP - interval poverenja;

AUC vrednosti postignute za parametar *Broj čestica* su se menjale od 0,50 (*nedovoljno dobar*, Prag 250) do 0,31 (*dobar*, Prag 200). Za Prag 170 je postignuta

još jedna prognostički značajna vrednost AUC = 0,38 koja je ujedno i jedina statistički značajna AUC vrednost postignuta za ovaj parametar.

AUC vrednosti postignute za parametar *Ukupna površina* su se menjale od 0,45 (AutoBin), rangirane kao *nedovoljno dobra*, do 0,63 i 0,37 (Prag 250 i Prag 200, respektivno) rangiranih kao *dobre*. Nijedna postignuta AUC vrednost *Ukupne površine* nije dostigla statistički značaj (na nivou značajnosti $p \leq 0,05$).

Parametar *Prosečna veličina* postiže dve prognostički značajne AUC vrednosti pri čemu obe nadmašuju maksimalnu AUC vrednost postignutu za kliničko-patološke parametre. Maksimalna i statistički značajna vrednost AUC = 0,74 rangirana kao *vrlo dobara* je ostvarena za Prag 250 dok je AUC = 0,67 postignut za Prag 240.



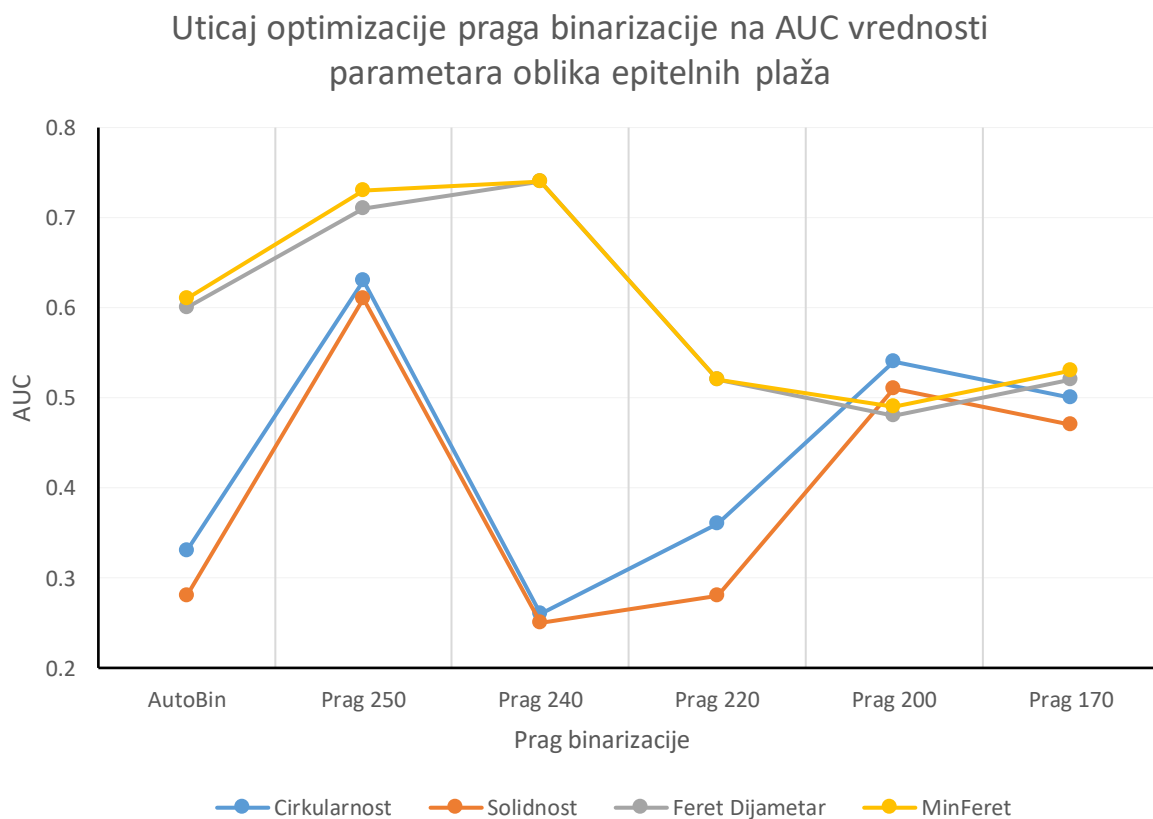
Slika 5.5.1. Prognostički učinak broja i parametara veličine epitelnih malignih plaža u funkciji praga binarizacije. Učinak je analiziran i kvantifikovan primenom ROC-AUC analize.

Parametar *Cirkularnost* je za četiri praga binarizacije ostvario prognostički značajan učinak (slika 5.5.2) pri čemu dve statistički značajne postignute AUC

vrednosti premašuju maksimalnu AUC vrednost postignutu za kliničko-patološke parametre: 0,33 (*dobar*, AutoBin) i 0,26 (*vrlo dobar*, Prag 240).

Za parametar *Solidnost* su za tri praga binarizacije, AutoBin, Prag 240 i 220, postignute prognostički značajne AUC vrednosti: 0,28, 0,25 i 0,28, respektivno. Sve tri postignute vrednosti su statistički značajne (na nivou značajnosti $p \leq 0,05$), rangirane su kao *vrlo dobre* i nadmašuju maksimalnu AUC vrednost postignutu za kliničko-patološke parametre.

Parametri *Feret dijametar* i *MinFeret* su za dva praga, Prag 240 i Prag 250, ostvarili prognostički značajan učinak rangiran kao *vrlo dobar*. Pri tome su postignute sledeće AUC vrednosti: 0,71 i 0,74 za *Feret dijametar*, i 0,73 i 0,74 za *MinFeret*.



Slika 5.5.2. Prognostički učinak parametara oblika epitelnih malignih plaža u funkciji praga binarizacije. Učinak je analiziran i kvantifikovan primenom ROC-AUC analize.

Svi analizirani parametri su za najmanje dva praga binarizacije ostvarili prognostički značajan učinak pri čemu su na slikama binarizovanim automatskom metodom samo parametri oblika postigli prognostički značajne AUC vrednosti.

Najmanje efikasni parametri su *Broj čestica* i *Prosečna veličina* koji su za samo dva praga binarizacije ostvarili prognostički značajan učinak, pri čemu je *Prosečna veličina* za Prag 250 ostvarila statistički (na nivou značajnosti $p \leq 0,05$) i prognostički značajnu AUC vrednost od 0,74 rangiranu kao *vrlo dobra*. Najuspešniji parametri su *Cirkularnost* i *Solidnost* koji su po četiri puta ostvarili prognostički značajan učinak. Pri tome je *Cirkularnost* postigla dve a *Solidnost* tri AUC vrednosti koje nadmašuju maksimalnu AUC vrednost postignutu za kliničko-patološke parametre.

5.5.2. Poboljšanje prognostičkog učinka morfometrijskih parametara usled optimizacije praga binarizacije i procena prognostički najefikasnijeg praga binarizacije

Optimizacija praga binarizacije je rezultovala poboljšanjem prognostičkog učinka svih analiziranih morfometrijskih parametara pri čemu su se kao najosetljiviji pokazali *Cirkularnost* i *Solidnost* sa maksimalnim povećanjem od $\Delta AUC = 0,24$ (tabela 5.5.2) ostvarenim na pragovanim slikama. Izuzetno poboljšanje u prognostičkom učinku i napredak od dve kategorije ($\Delta AUC \geq 0,2$) je ostvaren za sva četiri parametra oblika (*Cirkularnost*, *Solidnost*, *Feret dijametar* i *MinFeret*) i kod parametra *Veličina*. Parametri oblika su se pokazali nešto osetljivijim od parametara veličine pri čemu su ostvarena prosečna povećanja od $\Delta AUC = 0,233$ i $\Delta AUC = 0,167$, respektivno.

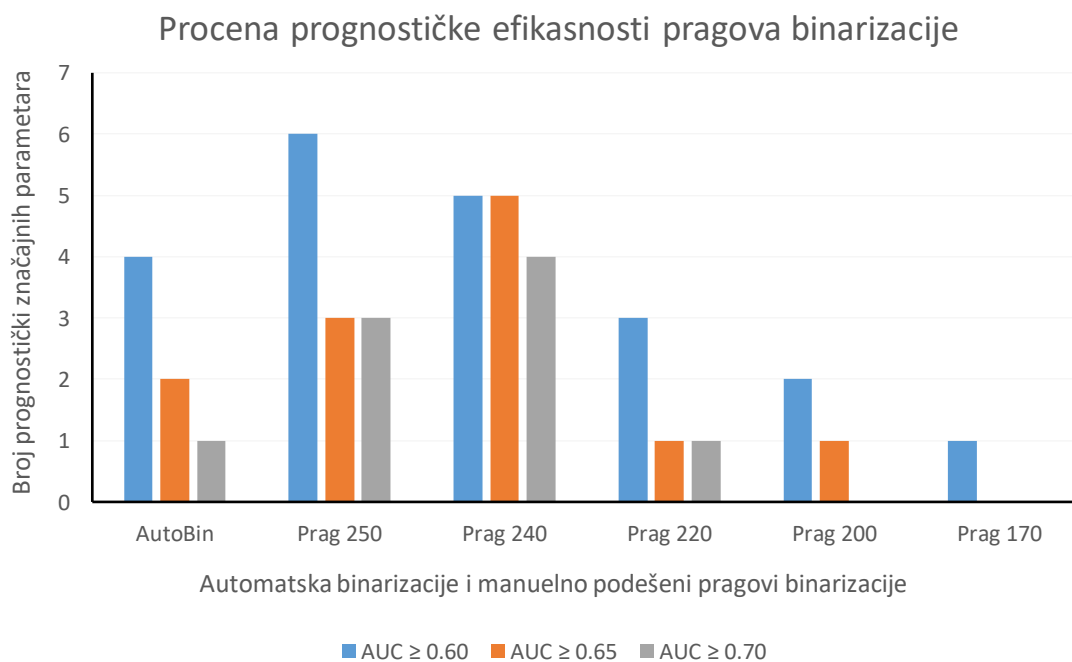
Tabela 5.5.2 Efekat promene praga binarizacije na prognostički učinak morfometrijskih parametara

	Broj čestica	Ukupna površina	Prosečna veličina	Cirkularnost	Solidnost	Feret dijametar	MinFeret
ΔAUC	0.19	0.08	0.23	0.24	0.24	0.22	0.23
ΔAUC^*	0.15	0.08	0.22	0.07	0.03	0.14	0.13

$\Delta AUC = AUC_{max} - AUC_{min}$ – doprinos ostvaren na pragovanim slikama; ΔAUC^* - maksimalni doprinos ostvaren na pragovanim slikama u odnosu na automatski binarizovane slike. Skraćenice: AUC - površina ispod ROC kriva;

Drugi značajan efekat optimizacije praga binarizacije, takođe ostvaren za sve morfometrijske parametre, je napredak u prognostičkom učinku na pragovanim slikama u odnosu na učinak ostvaren na automatski binarizovanim slikama. Najveće poboljšanje je zabeleženo za parametar *Prosečna veličina* koji je napredova od *nedovoljno dobrog* (AUC = 0,52, AutoBin) do *vrlo dobrog* (AUC = 0,74, Prag 250). Za parametar *Broj čestica* je zabeležen napredak od *nedovoljno dobar* u *dobar*, dok su *Feret dijametar* i *MinFeret* unapređeni od *dobar* u *vrlo dobar*.

U funkciji broja parametara koji su ostvarili prognostički značajan učinak najmanje je uspešan Prag 170 na kome je samo parametar *Broj čestica* ostvario značajan učinak sa AUC = 0,38. Učinak svih ostalih parametara na ovom pragu je rangiran kao *nedovoljno dobar*. Najviše značajnih parametara je identifikovano za Prag 250, čak šest, sa prosečnom AUC vrednošću od 0,68 (slika 5.5.3). Na ovom pragu tri parametra *vrlo dobrog* učinka su postignutim AUC vrednostima nadmašila maksimalnu AUC vrednost postignutu za kliničko-patološke parametre. *Najuspešniji prag binarizacije* je Prag 240 na kome je pet parametara postiglo AUC vrednosti veće od maksimalne AUC vrednosti postignute za kliničko-patološke parametre pri čemu su četiri rangiran kao *vrlo dobre* a ostvarena prosečna AUC vrednost e je 0,73.



Slika 5.5.3. Procena prognostički najefikasnijeg praga binarizacije. Procena je izvršena ustanovljavanjem i poređenjem broja parametara koji su postigli prognostički značajan učinak (AUC ≥ 0,60) pri automatskoj binarizaciji i za korisnički podešene pragove.

5.5.3. Prognostički učinak morfometrijskih parametara pri optimalnoj kombinaciji vrednosti praga binarizacije i filtera veličine i cirkularnosti

U drugoj fazi analize prognostičkog učinka morfometrijskih parametara epitelnih malignih plaža uz varijaciju praga binarizacije (šest pragova) primenjeni su i filteri *cirkularnosti* (pet filtera) i *veličine* (četiri filtera). Vrednosti donjeg praga filtera *cirkularnosti* koje su uzete u razmatranje su 0,0, 0,2, 0,4, 0,6 i 0,8 dok je maksimalna vrednost fiksirana na 1,0. Maksimalna vrednost filtera *veličine* je postavljena na beskonačno dok su kao donji prag razmatrane veličine 10, 20, 50 i 100. Kombinacijom odabranih pragova filtriranja za svaki prag binarizacije je dobijeno 20 setova slika.

Tabela 5.5.3. Prognostički učinak morfometrijskih parametara dobijenim pri optimalnoj kombinaciji filtera cirkularnosti (0,8 - 1,0) i veličine (20 - beskonačno) za prag binarizacije 250

Parametar	AUC ^a	95% IP	<i>p</i> - vrednost
Broj čestica	0,82	0,72-0,90	<0,001*
Ukupna površina	0,77	0,68-0,86	<0,001*
Prosečna veličina	0,63	0,54-0,77	0,03*
Cirkularnost	0,62	0,51-0,74	0,09
Solidnost	0,50	0,37-0,62	0,95
Feret dijametar	0,63	0,49-0,76	0,08
MinFeret	0,68	0,56-0,78	0,06

^aIzvršena "bootstrap" interna validacija; **p*≤0,05; Skraćenice: AUC - površina ispod ROC krive; IP - interval poverenja;

Pregledom rezultata dobijenih analizom svih 120 setova slika ustanovljeno je da optimalna kombinacija filtera pri kojoj morfolometrijski parametri ostvaruju najbolji prognostički učinak i postižu najviše AUC vrednosti (tabela 5.5.3) odgovaraju vrednosti: *prag binarizacije 250, interval filtera cirkularnosti 0,8 - 1,0 i interval filtera veličine 20 - beskonačno.*

Za optimalnu kombinaciju filtera je identifikovano šest morfolometrijskih parametra koji su ostvarili prognostički značajan učinak od čega su tri postigli AUC vrednosti koje nadmašuju maksimalnu AUC vrednost postignutu za kliničko-patološke parametre. Najbolji prognostički učinak je ostvaren za parametar *Broj čestica*. Rangiran je kao *odličan* uz postizanje statistički značajne (na nivou značajnosti $p \leq 0,05$) AUC vrednost od 0,82. Visoka AUC vrednost dobijena za *Broj čestica* je konzistentna sa frekventnom pojavom metastaza kod pacijenta sa visokim brojem raspršenih individualnih epitelnih čestica (tabela 5.1.1). Preostala dva parametra koja su nadmašili maksimalnu AUC vrednost postignutu za kliničko-patološke parametre su *Ukupna površina* koja je rangirana kao *vrlo dobra* (AUC = 0,77) i *MinFerret* koji je rangiran kao *dobar* (AUC = 0,68). Parametri *Prosečna veličina* (AUC = 0,63), *Cirkularnost* (AUC = 0,62) i *Feret dijametar* (AUC = 0,63) su rangirani kao *dobri* ali sa postignutim AUC vrednostima ispod maksimalne AUC vrednosti postignute za kliničko-patološke parametre. Parametar *Solidnost* je pri optimalnoj kombinaciji filtera izgubio prognostički značaj u potpunosti (AUC = 0,50).

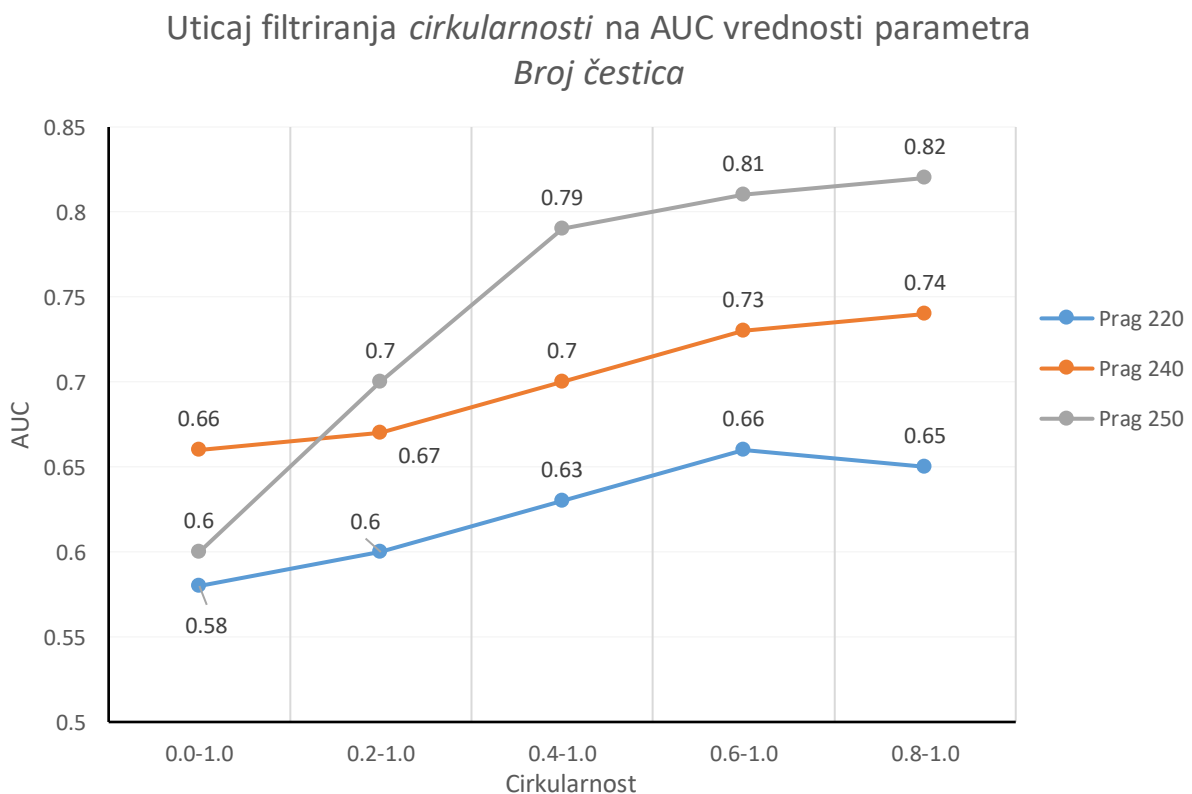
Ilustracija efekta filtracije *cirkularnosti* i *veličine* epitelnih malignih plaža na prognostički učinak parametra *Broj čestica* za najefikasnije pragove binarizacije (Prag 250, 240 i 220) je prikazan na slikama 5.5.4 i 5.5.5.

Za čestice veće od 20 piksela povećanje donjeg praga filtera *cirkularnosti* je rezultovalo poboljšanje AUC vrednosti za sva tri praga binarizacije (slika 5.5.4). Za Prag 250 je ostvareno maksimalno povećanje od $\Delta AUC = 0,22$ pri čemu je AUC porastao od 0,60 do 0,82. Usled ostvarenog poboljšanja prognostički učinak *Broja čestica* je napredovao za dva ranga, od *dobar* do *odličan*. Ostvareno povećanje za AUC vrednosti *Broja čestica* za Prag 240 je $\Delta AUC = 0,08$ sa minimalnom i maksimalnom postignutom AUC vrednošću od 0,66 (*dobar*) i 0,74 (*vrlo dobar*). Identično

poboljšanje učinka je ostvareno i za Prag 220 za koji su se AUC vrednosti se menjale od 0,58 (*nedovoljno dobar*) do 0,66 (*dobar*).

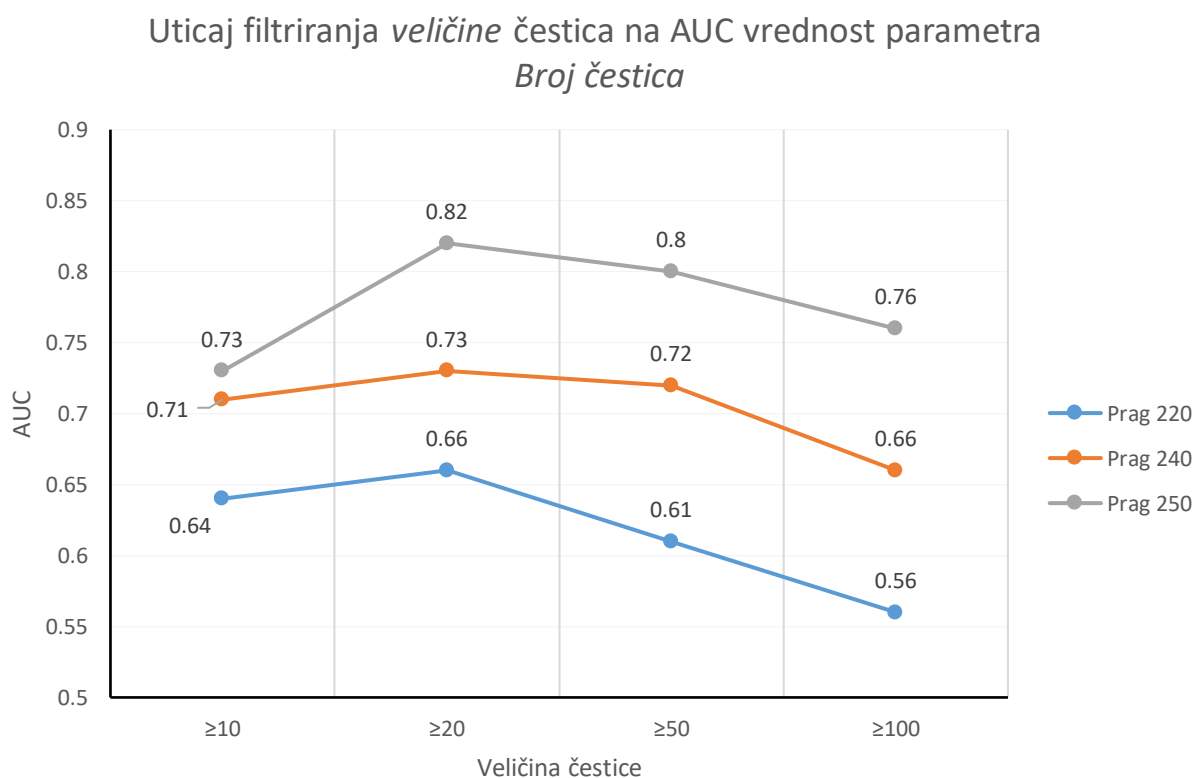
Za donji prag *cirkularnosti* postavljen na 0,2 postignute AUC vrednost su konzistentno veće za viši prag binarizacije. Usled toga su AUC vrednosti postignute za Prag 250 veće u odnosu za AUC vrednosti postignute za Prag 240 i 220, a AUC vrednosti postignute za Prag 240 su veće od vrednosti postignutog za Prag 220.

Za Prag 220 parametar *Broj čestica* samo za interval cirkularnosti 0,6 - 1,0 ostvaruje prognostički učinak koji nadmašuje učinak standardno korišćenih parametra uz postignutu AUC vrednost od 0,66. Za Prag 240 sve postignute AUC vrednosti za *Broj čestica* su veće 0,65 bez obzira na oblik, tj. bez obzira na vrednosti filtera *cirkularnosti*. Za čestice veličine veće od 20 parametar *Broj čestica* ostvaruje najefikasniji prognostički učinak za Prag 250 pri čemu su dve postignute AUC vrednosti rangirane kao *vrlo dobre* a dve kao *odlične* (slika 5.5.4).



Slika 5.5.4. Prognostički učinak parametra *Broj čestica* u funkciji filtera *cirkularnosti* za najefikasnije pragove binarizacije; filter *veličina* je fiksiran na 20 - beskonačno.

Za čestice čija je *cirkularnost* $\geq 0,80$ a *veličina* ≥ 20 za tri najefikasnija praga binarizacije filtriranje *veličine* je imalo obrnut efekat od efekta filtriranja *cirkularnosti* (slika 5.5.5). Usled toga je za parametar *Broj čestica* povećanje donjeg praga filtera *veličine* rezultovalo opadanjem prognostičkog učinka. Najveći gubitak učinka je zabeležen za Prag 220 sa smanjenjem od $\Delta AUC = -0,10$ pri čemu su se postignute AUC vrednosti menjale od 0,66 do 0,56. Za dati podskup čestica (*cirkularnost* $\geq 0,80$, *veličina* $\geq 0,20$) najmanji efekat varijacije *veličine* čestica na prognostički učinak *Broja čestica* je zabeležen za Prag 250 sa postignutim AUC vrednostima koje su opadale od 0,82 do 0,76 i zabeleženim smanjenjem od $\Delta AUC = -0,06$

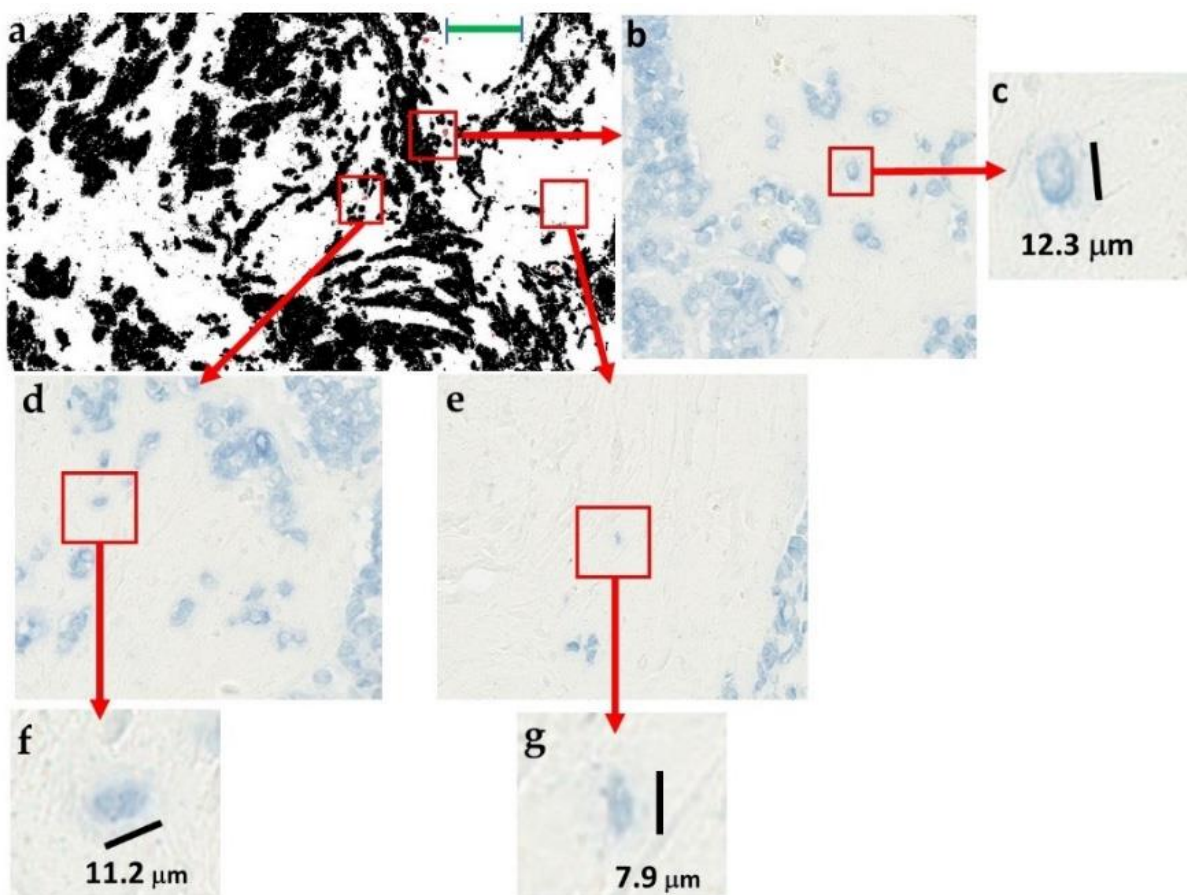


Slika 5.5.5. Prognostički učinak parametra *Broj čestica* u funkciji filtera *veličine* za najefikasnije pragove binarizacije; filter *cirkularnost* je fiksiran na 0,8 - 1,0.

5.5.4. Identifikacija prognostički najznačajnijih čestica

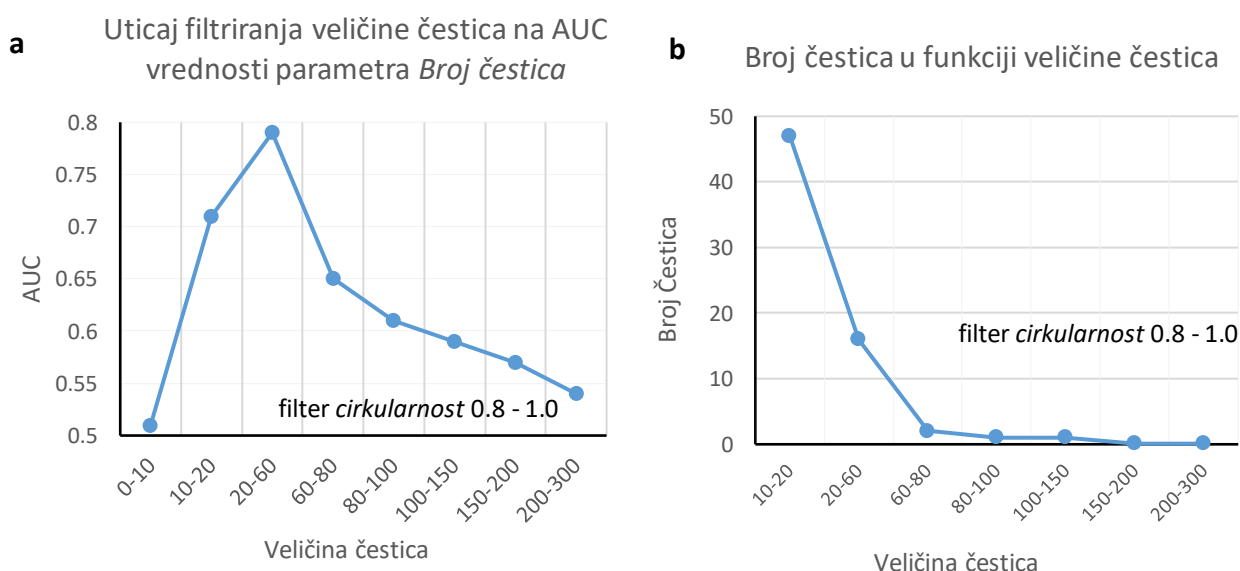
Identifikacija prognostički najznačajnijih čestica na optimalnom setu slika je bazirana na podudaranju čestica odgovarajućih karakteristika (*cirkularnost* 0.8 - 1.0 i *veličina*

20 - beskonačno) identifikovanih na binarnim slikama (slika 5.5.2 a) sa identičnim česticama na originalnim plavim slikama (slika 5.5.2 b, d, e). Prostorna raspodela odabranih čestica se može posmatrati i analizirati na slikama dobijenim preklapanjem originalne slike sa binarnom maskom na kojoj su odgovarajuće čestice oivičene crvenom konturom (slika 5.5.6. a). Na uvećanim segmentima koji sadrže odabrane čestice i prikazani su u originalnoj plavoj boji (slike 5.5.6. b, d, e) se mogu videti morfološki detalji čestica i njihova stromalna lokalizacija. Vizuelna analiza odabranih čestica na maksimalnom uvećanju (slike 5.5.6. c, f, g) i njihova pan-citokeratin imunoreaktivnost zajedno sa veličinom i približno kružnim oblicima su indikatori koji upućuju na to da se radi o pojedinačnim epitelnim ćelijama rasutim unutar strome između većih epitelnih malignih plaža.



Slika 5.5.6. Identifikacija prognostički najznačajnijih epitelnih malignih plaža u histopatološkim uzorcima raka dojke bojenih pan-citokeratinom. (a) Binarna slika uzorka na kojoj su čestice *cirkularnosti* ≥ 0.8 i *veličine* ≥ 20 obojene u crveno. (b, d, e) Uvećani segmenti plavih slika prikazuju odabrane čestice rasute po stromi i odvojene od većih epitelnih plaža. (c, f, g) Maksimalno uvećane odabrane čestice identifikovane kao pojedinačne epitelne maligne ćelije. Skala za (a): 300 μm

Identifikacija rasutih individualnih epitelnih ćelija kao prognostički najznačajnijih čestica je podržana i analizom prognostičkog učinka parametra *Broj čestica* i broja čestica u funkciji segmentacije čestica po njihovoj veličini (slika 5.5.7. a, b) pri fiksnom filteru cirkularnosti podešenim na 0.8 - 1.0. Najefikasniji prognostički učinak (AUC = 0.80) je ostvaren upravo za čestice veličine iz opsega 20 - 60 piksela čiji se dijametar aproksimativno menjao u opsegu 7-11 μm što odgovara dijametru individualnih ćelija. Takođe, broj čestica čija veličina prevazilazi veličinu individualnih epitelnih ćelija je bio veoma mali (slika 5.5.7 b). Broj čestica manjih od 20 piksela je bio veći ali su one bez prognostičkog značaja (slika 5.5.7 a) jer su postignute AUC vrednosti vrlo bliske dijagonali.



Slika 5.5.7. (a) Prognostički učinak parametra *Broj čestica* u funkciji segmentacije veličine čestica. Maksimalna postignuta AUC vrednost od 0.80 je zabeležena za čestice veličine iz intervala 20 - 60 piksela (dijametra 7 - 11 μm , aproksimativno). (b) Distribucija broja čestica u funkciji segmentacije veličine čestica. Manje čestice su brojnije; u proseku je nađeno 47 čestica čije je veličina u intervalu 10 - 20 piksela, 16 čestica u intervalu veličine 20 - 60 piksela, a samo četiri su bile veće od 60 piksela.

5.6. Multivarijaciona Cox-ova analiza

Rezultati Cox-ove analize nezavisnosti i intenziteta efekta teksturalnih i morfometrijskih parametara na pojavu metastaza su prikazani u tabelama 5.6.1, 5.6.2 i 5.6.3. Kriterijum za odabir parametara i njihovo uključivanje u analizu je bio statistički značaj postignutih AUC vrednosti; uzeti su u obzir parametri koji su zadovoljili uslov $p \leq 0,05$. Intenzitet efekta je kvantifikovan korišćenjem stope rizika HR pri čemu je $HR < 1$ indikator smanjenog a $HR > 1$ indikator povećanog rizika za pojavu metastaza.

Rezultati multivarijacione binarne logističke regresione analize starosti, kliničko-patoloških i teksturalnih parametara izvršenih na 8-bitnim slikama i u podintervalima su prikazani u tabeli 5.6.1. Da bi se izbegla pristrasnost usled kategorizacije, u analizi su korišćene kontinuirane vrednosti parametara. Identifikacija podskupa nezavisnih i statistički značajnih parametara je izvršena u procesu postupne selekcije pri čemu je kriterijum zadržavanja bio $p \leq 0,05$. Od standardno korišćenih parametara, statističku značajnost je zadržao samo parametar *Veličina tumora*, dok je u podintervalima kao statistički značajan parametar preostao samo *Koeficijent spljoštenosti*.

Tabela 5.6.1. Procena veličine efekta i nezavisnosti kliničko-patoloških i teksturalnih parametara u celokupnom intervalu intenziteta i u podintervalima

Parametar	p - vrednost	HR ^a	95% IP
Veličina tumora	0,03*	1,07	1,01-1,15
Srednja vrednost (interval 0-255)	0,003*	0,85	0,76-0,95
Entropija (interval 0-255)	0,03*	35,0	1,29-944
Koeficijent spljoštenosti (interval 240-255)	0,002*	1,02	1,01-1,03

^aIzvršena "bootstrap" interna validacija; * $p \leq 0,05$ Skraćenice: HR – stopa rizika, IP – interval poverenja.

U tabeli 5.6.2. su prikazani rezultati univarijacione analize statistički značajnih morfometrijskih parametara računatih na optimalnom setu slika.

Tabela 5.6.2. Procena veličine efekta morfometrijskih parametara računatih na optimalnom setu slika^a

Parametar	<i>p</i> - vrednost	HR ^b	95% IP
Broj čestica	0,001*	14,8	5,3 – 242
Ukupna površina	0,001*	17,2	5,2 – 230
Prosečna veličina	0,001*	11,9	2,1 – 28,8
Cirkularnost	0,02*	12,8	3,6 – 61,9

^aOptimalni set slika: cirkularnost 0,8 - 10, veličina 20 - beskonačno, prag binarizacije 250. ^bIzvršena "bootstrap" interna validacija; **p* ≤ 0,05 Skraćenice: HR – stopa rizika, IP – interval poverenja.

Rezultati multivarijacione analize koja je uključila kliničko-patološke i morfometrijske parametre računane na optimalnom setu slika su prikazani u tabeli 5.6.3. Kriterijum za odabir parametara i njihovo uključivanje u analizu je bio statistički značaj postignutih AUC vrednosti; uzeti su u obzir parametri koji su zadovoljili uslov $p \leq 0,2$. Identifikacija podskupa nezavisnih i statistički značajnih parametara je izvršena u procesu postupne selekcije pri čemu je kriterijum zadržavanja bio $p \leq 0,05$.

Tabela 5.6.3. Procena veličine efekta i nezavisnosti kliničko-patoloških i morfometrijskih parametara računatih na optimalnom setu slika^a

Parametar	<i>p</i> - vrednost	HR ^b	95% IP
Starost	0,03*	5,5	1,4-1264869
Broj čestica	0,001*	16,1	5,7-344551

^aOptimalni set slika: cirkularnost 0,8 - 10, veličina 20 - beskonačno, prag binarizacije 250. ^bIzvršena "bootstrap" interna validacija; **p* ≤ 0,05 Skraćenice: HR – stopa rizika, IP – interval poverenja.

6. DISKUSIJA

U grupi pacijentkinja analiziranih za potrebe ove teze rak dojke je dijagnostikivan 1993. godine a pacijentkinje su na osnovu procenjene patološke veličine tumora (pT1 i pT2) i ustanovljenog gradusa (G1) klasifikovane kao pacijenti niskog rizika. U skladu sa tada važećim protokolom pacijentkinje su tretirane samo lokalno: podvrgnute su operativnom zahvatu i radijaciji. Ovakav izbor pacijentkinja je glavna prednost analizirane grupe jer isključuje uticaj hormonskih ili citotoksičnih lekova na tok bolesti i pojavu metastaza.

Analiza prognostičke vrednosti starosti pacijenata i standardno korišćenih kliničko-patoloških parametra u analiziranoj grupi pacijentkinja je ukazala na njihov nizak učinak. Uzeto je u obzir osam parametara od kojih su samo tri rangirana kao *dobri* pri čemu se *Veličina Tumora* pokazala kao najefikasniji prognostički parametar sa postignutom AUC vrednošću od 0,65. Čak pet standardno korišćenih parametara su ocenjeni kao *nedovoljno dobri* sa postignutim AUC vrednostima bliskim 0,50. Inter i intratumorska heterogenost je vrlo izražena i jedna je od bitnih karakteristika raka dojke (33) koja je i glavni razlog za poteškoće u prognozi i individualizaciji lečenja. Niska pojedinačna efikasnost zahteva kombinovanu upotrebu izabranih parametara pa se tako zarad postizanja prognoze maksimalne preciznosti koriste protokoli koji kombinuju nekoliko dostupnih karakteristike i tumora i pacijenta (45).

Izbor pacijentkinja bez proširenja na limfne čvorove, metastaza i sa manjim tumorima i minimalnim histološkim gradusom ograničava korišćenje TNM i histološkog gradusa kao glavnih prognostičkih faktora u analiziranoj grupi pacijenata (123). Štaviše, u ovoj grupi pacijentkinja sa rano dijagnostikivanim karcinomom, malom veličinom tumora, bez zahvaćenih limfnih čvorova i bez metastaza, očigledna je potreba za novim prognostičkim parametrima.

6.1. Efekti segmentacije histograma na prognostički učinak teksturalnih parametara histopatoloških uzoraka tumora dojke

6.1.1. Ekspresija citokeratina u epitelnim malignim ćelijama

Citoskelet eukariotskih ćelija je izgrađen od više desetina proteina koji obrazuju složenu mrežastu strukturu koja se prostire kroz citoplazmu. Citoskelet ćeliji omogućava funkcije kao što su promena oblika, kretanje, deoba i komunikacija. Kombinacija biohemijskih analiza i mikroskopskih posmatranja je omogućila da se na osnovu razlika u kvalitativnim i kvantitativnim karakteristikama identifikuju tri osnovne komponente citoskeleta: mikrofilamenti (aktinski filamenti), mikrotubule i intermedijarni filamenti (124).

Intermedijarni filamenti su strukturno najkompleksniji, sastoje se od više stotina različitih elemenata i uobičajeno je da se dele u pet grupa. Prva i druga grupa intermedijarnih filamenata je izgrađena od kiselih i baznih citokeratina (125), respektivno. Prisutnost intermedijarnih filamenata je za sada sa sigurnošću potvrđena u životinjskim ćelijama kojima obezbeđuju mehaničku potporu.

Citokeratini se mogu naći u povećanim količinama u epitelnim ćelijama kod više tipova epitelnog kancera, na primer, karcinoma dojke i pluća (126). Do sada je identifikovano više od 20 citokeratina (127–129) različitih struktura i molekulskih masa. Klasifikovani su u dve grupe: tip II koga čine bazni citokeratini identifikovani brojevima od 1 do 8 i tip I koga čine kiseli citokeratini identifikovani brojevima od 9 do 20. Nađeno je da su u malignim ćelijama najviše prisutni citokeratini 8, 18 i 19 (130).

Ćelijski tip i stepen diferencijacije ćelija su glavni faktori koji određuju ekspresiju citokeratina. Više diferencirane epitelne ćelije imaju više citokeratina (131). Obrazac proizvodnje citokeratina u toku neoplastične transformacije je uglavnom malo podložan promenama što ga čini vrlo pogodnim za upotrebu kao tumor markera (132,133).

Imunohistohemijske metode analize citokeratina obojenog upotrebom koktela monoklonskih antitela omogućavaju analizu njegove ekspresije. Jedno od uobičajeno korišćenih antitela koja se vezuju za citokeratine je AE1/AE3 koktel antitela koji se najčešće koristi za diferencijaciju epitelnih od neepitelnih tumora (131).

Izuzetan prognostički potencijal pan-citokeratinskog imunobojenja je rezultat činjenice da takvo bojenje omogućava vizualizaciju i analizu raspodele, veličine i oblika malignih plaža nastalih neoplastičnom transformacijom epitelnih ćelija. Bojenje malignih plaža citokeratinom (60) i njihov značaj u prognozi karcinoma dojke je ranije ustanovljen (15,64,86,134,135). Za potrebe ove teze je prvi put izvršena sistematska studija pan-citokeratinskog imunobojenja bazirana na analizi raspodele intenziteta bojenja. Time je omogućeno je da se ustanovi značaj i kvantifikacija prognostičkog učinka parametara koji opisuju prostornu raspodelu intenziteta bojenja na 8-bitnim slikama sive skale i na nepreklapajućim podintervalima celokupnog opsega intenziteta (61).

6.1.2. Učinak u podintervalima niskih i srednjih intenziteta imunobojenja

Prethodna istraživanja pan-citokeratinskog imunobojenja su ukazala na povezanost niskog intenziteta bojenja sa povećanim rizikom pojave metastaza i na visok prognostički učinak parametra *Srednja Vrednost* intenziteta piksela (62). Treba imati u vidu da su intenzitet sivog tonaliteta (intenzitet piksela) i intenzitet bojenja (količina pan-citokeratina) obrnuto proporcionalni. Naime, kako je crnim pikselima pridružen intenzitet 0 a belim 255 proizilazi da tamniji pikseli imaju niži intenzitet iako odražavaju veću količinu pan-citokeratina i jače bojenje.

Razdvajanjem histograma na podintervale je napravljen korak dalje u odnosu na prethodne studije i detaljniji uvid u prognostičke performanse intenziteta pan-citokeratinskog imunobojenja. Analizirani teksturalni parametri su na svih sedam podintervala histograma ostvarili prognostički značajan učinak pri čemu je postignuto 36 prognostički značajnih AUC vrednosti. Od toga su tri AUC vrednosti bile na nivou maksimalne AUC vrednosti postignute za kliničko-patološke parametre a čak 17 postignutih AUC vrednosti su bile veće od 0,65 pri čemu su tri ocenjene kao *vrlo dobre*

sa $AUC \geq 0,70$. Ovako dobri prognostički rezultati su ukazali da i podintervali intenziteta pan-citokeratinskog imunobojenja obezbeđuju dovoljno informacija o heterogenosti tumora i da se korišćenjem teksturalnih parametara na ovim podintervalima može postići klasifikacija pacijenata efikasnija od klasifikacije postignute korišćenjem standardnih demografskih i kliničko-patoloških parametara.

Prognostički značajan učinak teksturalnih parametara je ostvaren i na 8-bitnim slikama i u podintervalima celokupnog intervala intenziteta. Najslabiji učinak je zabeležen u podintervalu tamnijih piksela intenziteta 0 - 130 i u podintervalu svetlih piksela visokih intenziteta 200 - 220. U podintervalu 200 - 220 su samo dva parametra ostvarila značajan učinak dok je u podintervalu najtamnijih piksela 0 - 130 čak šest parametara ostvarilo značajan učinak. Iznenadujuće je da su svih osam prognostički značajnih AUC vrednosti postignutih na ovim podintervalima bez statističkog značaja (na nivou značajnosti $p \leq 0,05$) a uz to su i niže su od maksimalne AUC vrednosti postignute za kliničko-patološke parametre.

Neoplastična transformacije epitelnih ćelija je praćena promenama u ekspresiji citokeratina kao posledica promena u citoskeletnoj strukturi. Ranije je ukazano na značajnu korelaciju između gradacije tumora sa ekspresijom citokeratina 8, 18 i 19 (136). Takođe je ustanovljeno i da viši intenzitet imunobojenja odgovara bolje diferenciranim epitelnim ćelijama (131). U svetlu ovih rezultata slab prognostički učinak teksturalnih parametara ekstrahovanih sa intervala tamnih piksela ukazuje na odsustvo prognostičkog značaja diferenciranih epitelnih ćelija kojima odgovara visoka ekspresija citokeratina. Razlog slabog učinka u podintervalu 200 - 220 je verovatno činjenica da pikseli iz ovog opsega odgovaraju granici koja razdvaja tumor od strome i zato ne sadrže značajne informacije o heterogenosti tumora koje bi mogle da se analiziraju teksturalnim parametrima.

Najefikasniji prognostički učinak je ostvaren u tri podintervala srednjih intenziteta piksela: 130 - 160, 160 - 180 i 180 - 200, na kojima su svih pet GLCM parametara kao i dva parametra statistike prvog reda ostvarili značajan učinak. Pri tome, u podintervalu 130 - 180 svi GLCM parametri po postignutim AUC vrednostima nadmašuju maksimalnu AUC vrednost postignutu za kliničko-patološke parametre. Ovako dobri rezultati GLCM parametara ukazuju na izražen prognostički značaj epitelnih ćelija umerenog intenziteta bojenja kao i na to da su značajne informacije o

heterogenosti tumora sadržane upravo u subintervalu srednjih intenziteta. Imajući u vidu da korišćeni GLCM parametri kvantifikuju heterogenost teksture, njihov izuzetan prognostički učinak ostvaren na subintervalu srednjih piksela je indikator bitnog doprinosa srednje diferenciranih ćelija heterogenosti tumora kao izvoru najznačajnijih informacija za efikasnu prognozu.

Dobar prognostički učinak GLCM parametara u podintervalu srednjih intenziteta 130 - 180 je podržan i *dobrim* učinkom parametra *Površina* koji je takođe ostvario značajan prognostički učinak na ovom podintervalu i čije su niske vrednosti povezane sa lošijim ishodom bolesti (61). Ovaj rezultat je donekle iznenađujući u skladu je sa prethodnim studijama prognostičke efikasnosti TSR indeksa (engl. *tumor stroma ratio*) kojim se iskazuje odnos između površine tumora i površine strome. Pokazano je da su vrednosti $TSR < 1$ povezane sa kraćim vremenom preživljavanja (137). Shodno tome, manje površine malignih plaža u tumorima ukazuju na lošiji ishod za kancer dojke kao i na lošiji ishod kod drugih vrsta raka (138,139). Isti argumenti se mogu prihvatiti i za objašnjenje prognostičke efikasnosti parametra *Površina* u intervalu najsvetlijih piksela 220 - 255 nespecifičnog bojenja koje odgovaraju stromi gde ovaj parametar postiže visoke vrednosti a učinak mu je takođe rangiran kao *dobar*. Ovi rezultati se mogu tumačiti u skladu sa prethodno ustanovljenim indirektnim efektom većih malignih oblasti u smanjenju veličine strome koja podstiče razvoj protumorskih imunoloških infiltrata (140). Takođe je ustanovljeno da je visok odnos površina tumor/stroma povezan sa niskim peritumorskim inflamatornim infiltratima kod pacijenata sa kolorektalnim karcinomom (141).

Analiza rezultata je ukazala na relativno veliku varijaciju u intenzitetu imunobojenja pan-citokeratinom ne samo unutar pojedinačnih malignih plaža, već i između pojedinačnih epitelnih ćelija. Opravdano je pretpostaviti da su varijacije u intenzitetu bojenja epitelnih ćelija posledica razlike u sadržaju citokeratina i razlika u ekspresiji ćelijskih pod tipova. Izražene varijacije intenziteta imuno bojenja potvrđuju i ističu intratumorsku heterogenost malignih klonova ukazujući na povezanost porasta heterogenosti sa lošijim ishodom bolesti što je u skladu sa prethodnim istraživanjima koja su ukazali na vezu između intratumorske heterogenosti bazirana na HER2 ekspresiji sa lošijim ishodom bolesti (142).

Uočene visoke varijacije u intenzitetu pan-citokeratinskog imunobojenja su rezultat fenotipske različitosti tumorskih ćelija kao posledica genetske poliklonalnosti susednih ćelija, ali je verovatnije da su najvećim delom posledica epigenetskih promena i stohastičkih procesa (143).

Intratumorska heterogenost pojedinačnih ćelija je prethodno istraživana kod raka dojke pri čemu je ukazano da povećana heterogenost ukazuje na lošiju prognozu (143,144). Istraživanja rađena za potrebe ove teze su po prvi put ukazala da varijacija intenziteta pan-citokeratinskog imunobojenja takođe može biti upotrebljena kao mera intratumorske i inter-ćelijske heterogenosti kao i da se može koristiti za dobijanje korisnih prognostičkih markera.

Procena intratumorske heterogenosti je od izuzetnog značaja iz više razloga, pre svega zbog svog prognostičkog značaja i jasne povezanosti sa agresivnošću bolesti i lošom prognozom, kao i zbog činjenice da je jedna od glavnih smetnji za efikasnu terapiju raka (143).

6.1.3. Učinak u podintervalima specifičnog i nespecifičnog imunobojenja

Podela histograma je omogućilo vizualizaciju epitelnih malignih plaža kojima odgovaraju različiti intenziteti pan-citokeratinskog imunobojenja. Vredi još jednom naglasiti da višim vrednostima intenziteta piksela odgovaraju svetliji pikseli i niži intenziteti bojenja. Najintenzivnije bojenje kojem odgovaraju najtamniji pikseli iz podintervala 0 - 130 je doprinelo vizualizaciji tek nekoliko ćelija u okviru malignih plaža. Za razliku od tamnih, pikseli srednjih intenziteta iz podintervala 130 - 200 su približno tri puta brojniji i rasuti su po celoj površini malignih plaža. Još svetliji pikseli iz intervala intenziteta 200 - 220 uglavnom odgovaraju ivicama malignih plaža dok su najsvetliji pikseli iz intervala 220 - 255 kojima odgovara najniži intenzitet bojenja uočeni samo u stromi.

Kako se citokeratin eksprimira samo u epitelnim ćelijama unutar tumorskih plaža a samo se sporadično uočava u stromi (15,145), može se pretpostaviti da pikseli intenziteta 220 približno markiraju granicu između specifičnog (0 - 220) i nespecifičnog bojenja (220 - 250), tj. definiše granicu između površina koje odgovaraju epitelnim malignim plažama i stromi, respektivno. Opravdano je predložiti

dodatna istraživanja kojima bi se potvrdila činjenica da pikseli intenziteta 220 omogućavaju precizno razdvajanje strome i tumora jer bi to moglo da poboljša preciznost izračunavanja TSR indeksa za koji je ustanovljeno da se može koristiti kao nezavisan prognostički parametar kod različitih tipova kancera (137). Trenutno se TSR indeks najčešće procenjuje mikroskopskim metodama analizom hematoksilin-eozinom (HE) bojenih uzoraka koji pokazuju ograničenja u računarskoj analizi usled slabog kontrasta između strome i tumora.

Značajno je i opažanje da prognostički učinak teksturalnih parametara u intervalu specifičnog bojenja nije nadmašio njihov učinak ostvaren u podintervalima. Evidentno je i da su originalne 8-bitne slike čiji histogram uključuje kako specifično tako i nespecifično bojenje omogućile postizanje AUC vrednosti koje nadmašuju vrednosti dobijene u podintervalima specifičnog pan-citokeratin bojenja. Najefikasniji prognostički učinak na 8-bitnim slikama je ostvaren za GLCM parametar ASM. Učinak ASM je rangiran kao *vrlo dobar* a postignuta AUC vrednost od 0,77 nadmašuje maksimalnu AUC vrednost postignutu za standardno korišćene parametre. Ovi nalazi su u skladu sa neočekivano niskim učinkom zabeleženim u intervalu specifičnog bojenja tamnijih piksela 0 - 130 kao i nešto boljem prognostičkom učinku dobijenim u intervalu nespecifičnog bojenja. Takođe, svi parametri statistike prvog reda su dostigli prognostički značaj u intervalu nespecifičnog bojenja kojem pripadaju najsvetliji pikseli. Zajedno, ovi rezultati ukazuju na prognostički značaj strome. Nova istraživanja ukazuju da biološke karakteristike tumora nisu jedini faktori koji utiču na njihovu progresiju i razvoj metastaza već se moraju uzeti u obzir i efekat interakcije između ćelija tumora i strome i drugih ćelijskih elemenata (146). Značaj strome se mora uzeti u obzir jer stroma promoviše progresiju tumora kroz produkciju hranljivih sastojaka i faktora rasta (137). U tom cilju, vidljiv je porast interesa za analizu osobina tumorskog okruženja uključujući i karakteristike strome (147).

Međusobna zavisnost i visoka negativna korelacija tumora i strome, tj. intervala specifičnog i nespecifičnog bojenja, je najverovatniji razlog inverzne prognostičke efikasnosti teksturalnih parametara u ovim podintervalima. Ovakav odnos je znatno izraženiji za GLCM parametre. Od parametra statistike prvog reda, inverzija AUC vrednosti se uočava samo za parametar *Površina*, koje se menjaju od $AUC < 0,50$ za interval 0 - 200 u $AUC > 0,50$ za interval 200 - 255, pri čemu se ostvaruje identičan

maksimalan prognostički učinak sa postignutim AUC vrednostima od 0,31 i 0,69. Istovetna tendencija je opažena za GLCM parametre CON i ENT pri čemu su parametri IDM i ASM pokazivali obrnutu tendenciju. To je bilo očekivano s obzirom da CON i ENT kvantifikuju heterogenost a IDM i ASM homogenost teksture tako da je njihov inverzni odnos očekivan.

Analiza histograma specifičnog pan-citokeratinskog imunobojenja je ukazala na dominaciju svetlijih piksela dok su tamniji pikseli iz intervala 0-160 bili znatno manje zastupljeni. Ovakva neravnomerna raspodela intenziteta je bila razlog da prvi podinterval bude znatno širi. Prvi interval koji odgovara najtamnijim pikselima sadrži 130 nivoa sive boje dok svetlijim intervalima odgovara između 15 i 20 sivih tonova. Specifično bojenje je, u proseku, pokrivalo 41% površine slike tako da slaba zastupljenost niskih intenziteta piksela nije posledica malog broja epitelnih ćelija. U prethodnim istraživanjima je pokazano da je pan-citokeratinsko bojenje zrelih epitelnih ćelija intenzivnije u poređenju sa manje diferenciranim ćelijama (131,148). Stoga se slaba prisutnost intenzivnog pan-citokeratinskog imunobojenja može razumeti kao posledica niske diferencijacije epitelnih ćelija u analiziranim uzorcima.

6.2. Efekti optimizacije praga binarizacije sive skale na prognostički učinak morfometrijskih parametara epitelnih malignih plaža

Za potrebe ove teze je prvi put izvršena opsežna analiza morfometrijskih karakteristika epitelnih malignih plaža na digitalnim slikama histopatoloških uzoraka raka dojke. Analiza plaža je izvršena u funkciji njihove veličine i oblika pri čemu je istraživanje bilo usmereno na identifikaciju novih prognostički značajnih morfometrijskih parametara. Istraživanja epitelnih ćelija je od posebnog značaja jer se većina karcinoma dojke razvija upravo neoplastičnom transformacijom ove vrste ćelija za koje je tipično da se grupišu u klastere izmenjene strukturne konzistencije u odnosu na normalno tkivo (149). Morfometrijske karakteristike epitelnih plaža su

povezane sa obrascima rasta tumora i njihove karakteristike mogu da se koriste kao prediktori ishoda bolesti (64,150). Na osnovu izražene varijabilnosti veličine i oblika bilo je opravdano pretpostaviti da kvantifikovane morfološke karakteristike epitelnih plaža mogu ostvariti značajan učinak u prognozi pojave metastaza. Imunohistohemijskim bojenjem pan-citokeratinom je omogućena jasna vizuelizacija plaža a zatim i relativno jednostavno izdvajanje od pozadine i analiza korišćenjem programa ImageJ. Računarska analiza izvršena za potrebe ove teze nije imala za cilj automatizaciju mikroskopske analize već je korišćena kao alat koji je omogućio sveobuhvatnu morfometrijsku analizu i klasifikaciju epitelnih plaža uočljivih u vidnom polju mikroskopa a čiji broj ide od nekoliko stotina do nekoliko hiljada klastera.

Prethodne studije koje su takođe za cilj imale identifikaciju novih prognostičkih parametara histopatoloških uzoraka uglavnom su bile bazirane na nespecifično bojenim uzorcima tkiva sa HE (4,20,22–24). Ovaj metod bojenja uzoraka je nisko specifičan i sa više boja u odnosu na imunobojenje epitelnih ćelija (150) usled čega su parametri dobijeni analizom HE bojenih uzoraka relativno kompleksni i njihova analiza zahteva upotrebu računarskih metoda (151). Nasuprot tome, analize uzoraka bojenih pan-citokeratinom su ukazale da je prognostički značajne pojedinačne epitelne ćelije rasute unutar strome jednostavno uočiti čak i standardnim mikroskopskim pregledom (15).

Wang et al. su u svojoj studiji ukazali na značaj osnovnih karakteristika epitelnih malignih plaža: broja, površine, obima i cirkularnosti, kao i da njihova kombinacija može da dovede do prognostički značajnih rezultata (123). To je bio motiv za proširenje analize prognostičke efikasnosti morfometrijskih karakteristika plaža u funkciji njihovog oblika i veličine kao i na ispitivanje efekta različitih metoda binarizacije slike.

U prethodnoj studiji histopatoloških uzoraka tumora dojke obojenih pan-citokeratinom Wang et al. su pokazali da je najefikasniji morfometrijski prognostički parametar *Cirkularnost* sa AUC vrednošću od 0,66 (123) rangirana kao *dobra*. Gotovo identičan rezultat je dobijen u analizi urađenoj na binarizovanim slikama procesom automatske binarizacije i bez primene morfometrijskih filtera *cirkularnosti* i *veličine*. Postignuta je AUC vrednost od 0,33 što je ekvivalentno AUC vrednosti 0,67, kao što je objašnjeno u poglavlju 4.6.1. Parametar *Cirkularnost* se pokazao osetljivim na

varijaciju praga binarizacije što je omogućilo da u odnosu na automatski binarizovane slike unapredi svoj prognostički učinak za jednu kategoriju uz poboljšanje od *dobrog* do *vrlo dobrog*. Sa pragom binarizacije 240, uzimajući u obzir čestice svih veličina i oblika, za *Cirkularnost* je postignuta AUC vrednost od 0,26.

Cirkularnost je morfometrijski parametar koji se koristi za kvantifikaciju oblika kombinovanjem geometrijskih i matematičkih osobina plaža a definisan je kao odnos površine i obima (92). Iako je definisan na jednostavan način, glavni nedostatak parametra *Cirkularnost* je činjenica da njegovo računanje zahteva primenu računara i specijalizovanih programa koja omogućavaju određivanje obima i površine objekata sadržanim u digitalnim slikama. Imajući u vidu činjenicu da računarske analize retko imaju za cilj i/ili retko uspevaju da precizno identifikuju jednostavne prognostičke karakteristike, može se zaključiti da je jedan od najupečatljivijih rezultata analize rađene za potrebe ove teze upravo identifikacija jednostavnih parametara i pogodnih za vizuelnu identifikaciju.

Prethodna istraživanja invazivnog potencijala i obrazaca grupisanja epitelnih malignih ćelija su sugerisala da je invazivnost tumora dojke više posledica karakteristika epitelnih plaža nego karakteristika pojedinačnih ćelija (17, 18). Takođe se smatra da izdanci tumora (engl. *tumour buddings*) raspolažu velikim prognostičkim potencijalom s obzirom da imaju karakteristike epitelno mezenhimalne tranzicije svojstvene za inicijalnu fazu i jednim od ključnih koraka metastaziranja raka dojke (145,152). Izdanci tumor se definišu kao mali klasteri jedna do pet nediferenciranih malignih ćelija, odvojeni od tumora i rasuti u stromi (153) skoro uvek u blizini invazivne granice tumora. Ovako definisani tumor izdanci uključuju i pojedinačne epitelne ćelije. Pored najčešćih peritumorskih izdanaka ovi mali klasteri malignih ćelija su opaženi i kao intratumorski (154). Kao faktor progresije i preživljavanja izdanci tumora su proučavani kod različitih tipova kancera uključujući i kancer dojke (155). Rezultati istraživanja upućuju na zaključak da tumori organizovani u manji broj malignih plaža ali većih dimenzija, imaju slabiji metastatski potencijal, verovatno usled smanjene mobilnosti, dok su tumori kod kojih je uočen veći broj izdanaka tumora povezani sa lošijim ishodom bolesti (156).

Varijacijom oblika i veličine epitelnih malignih plaža uz optimizaciju praga binarizacije omogućena je identifikacija prognostički najznačajnijih klastera. Na

slikama za koje je prag binarizacije fiksiran na 250 a donje vrednosti filtera cirkularnosti i veličine postavljeni na 0,8 i 20, respektivno, za parametar *Broj čestica* ostvaren je *odličan* prognostički učinak sa postignutom AUC vrednošću od 0,82 i stopom rizika HR = 16,1. Time su nadmašeni rezultati (AUC = 0,66 i HR = 1,45) koje su ostvarili Wang et al. (123) u svojoj studiji prognostičke efikasnosti matematičkih parametara dobijenih kompjuterskom analizom izdanaka tumora. Iznenađujući je rezultat da su strukture koje zadovoljavaju filtere veličine i oblika koji dovode do prognostički najuspešnijih rezultata identifikovane kao pojedinačne epitelne ćelije lokalizovane u stromi i rasute između većih klastera. Izuzetan prognostički uspeh ovih ćelija se može razumeti u svetlu činjenice da normalne epitelne ćelije u dojci formiraju klastere definisanih oblika u perenhimu tako da je identifikacija rasutih pojedinačnih epitelnih ćelija u stromi indikator povećanog maligniteta i lošijeg ishoda.

Prognostički uspeh rasutih individualnih epitelnih ćelija je u skladu sa rezultatima dobijenim analizom izdanaka tumora čije je prisustvo takođe povezano sa nepovoljnim kliničkim ishodom (145) uz razlike koje se odnose na veličinu, lokalizaciju i prognostički učinak. Najvažnija razlika je poboljšani prognostički učinak uz zabeleženi HR = 14,6 u odnosu na HR = 6,5 koji odgovara izdancima tumora (145). Osim toga, izdanci tumora su uglavnom veći od jedne ćelije i primarno su ispitivani u peritumorskom području za razliku od prognostički najefikasnijih pojedinačnih ćelija lokalizovanih unutar tumora.

Cirkulišuće tumorske ćelije (CTC⁹) su maligne ćelije koje su se odvojile od primarnog tumora i ušle u krvotok po kome se kreću u toku metastatskog širenja (157). Prisustvo CTC je potvrđeno kod oko 30% pacijenata sa dijagnostikovanom metastazama raka dojke (158). Ove ćelije su izazvale veliku pažnju (159) od kada je ukazano da je njihovo prisustvo povezano sa pojavama metastaza pri čemu su kao nezavisani prognostički faktor postigle AUC vrednost od 0,73 (160). Metode identifikacije CTC se nameću kao manje invazivna alternativa klasičnim tkivnim biopsijama. Analize rađena za potrebe ove teze su ukazale da je broj čestica identifikovanih u tumorskoj stromi, čak i pre njihove diseminacije u krvotok,

⁹ Akronim izraza na engleskom "Circulating Tumor Cells".

dragocen prognostički faktor pri čemu su AUC vrednosti postignute za parametar *Broj čestica* od 0,82 pri optimalnim vrednostima filtera cirkularnosti i veličine nadmašuju vrednosti postignute za CTC. Superiorni prognostički rezultati postignuti analizom epitelnih ćelija lokalizovanih u stromi u odnosu na prognostički učinak CTC bi se mogli objasniti kao posledica više faktora, kao na primer značajne molekulske i funkcionalne heterogenosti CTC (33) kao i niže preciznosti metoda koje se koriste u identifikaciji CTC (161). Nadalje, kao bitna prednost pristupa primenjenog u okviru ove teze se nameće jednostavnost i pouzdanost računarske kvantifikacije karakteristika uzoraka tkiva izvršene kompjuterskom analizom.

Svi morfometrijski parametri su se pokazali osetljivim na varijaciju praga binarizacije, pri čemu su parametri oblika osetljiviji od parametara veličine. Usled toga je varijacijom praga, mereno na AUC skali, prognostički učinak poboljšan čak za dve kategorije, od *nedovoljnog* do *vrlo dobrog*. Povećanje praga binarizacije je imalo pozitivan efekat na prognostički učinak što se može razumeti imajući u vidu činjenicu da slabo diferenciranim epitelnim ćelijama odgovara niska akumulacija citokeratina usled čega su na slikama predstavljene pikselima visokog intenziteta. Najefikasniji prognostički učinak je ostvaren na najvišim pragovima binarizacije, 240 i 250, kojima je omogućeno uzimanje u obzir najsvetlije obojenih, slabo diferenciranih ćelija koje su rasute po stromi a koje su se pokazale kao prognostički najznačajnije. Imajući u vidu veličinu pojedinačnih ćelija koja je približna 10 μm a kojima odgovara između 20 i 60 piksela, može se objasniti i pozitivan efekat filtriranja veličine čestica na prognostički učinak morfometrijskih faktora koji su beležili porast usled sniženja donje granice filtera veličine. Vrednost filtera veličine pri kojem je ostvarena najviša AUC vrednost je 20 piksela što je ukazalo na pojedinačne epitelne ćelije.

Jedan od limitirajućih faktora ove studije je njen retrospektivni dizajn. Prognostički modeli uzimaju u obzir faktore vrlo različite prirode koji predstavljaju karakteristike pacijenta ili su povezani sa karakteristikama bolesti. Iako mogu da budu međusobno različiti, prognostički modeli se razvijaju u cilju poboljšanja i standardizacije prognoze (162,163). Za razvoj prognostičkog modela se koriste uzorci pacijenta za koje je već ustanovljeno prisustvo odabranog ishoda i pretpostavljena je njegova veza sa kovarijantama od interesa (104). S obzirom da se prognostički modeli

razvijaju za primenu na pacijentima za koje ishod bolesti nije unapred poznat neophodno je izvršiti validaciju modela koja uključuje procenu generalizabilnosti i prognostičkog značaja modela procesima interne i eksterne validacije. Razlozi za preoptimističnost (engl. *overfitting*) modela može da bude veliki broj prognostičkih faktora i mali broj pacijenata (164,165). Iako ispitivana grupa od 102 pacijentkinje zadovoljava uslove uspostavljene prospektivnom analizom veličine uzorka, opravdano se može pretpostaviti da je veličina analiziranog uzorka jedan od ograničavajućih faktora ove studije.

Za kontrolu preoptimističnosti rezultata u ovoj tezi je izvršena interna validacija primenom bootstrap metode za rezultate ROC i COX-ove regresione analize. Bootstrap validacija je tehnika simulacije u kojoj se model testira na modifikovanoj grupi originalnih uzoraka pomoću nasumične zamene (120). U procesu odabira bootstrap uzorka postoji verovatnoća da neki pacijent bude izabran nekoliko puta a neki ni jednom. U cilju kontrole preoptimističnosti rezultata, za potrebe ove teze je izvedeno 1 000 bootstrap ponavljanja (15,61).

Za pouzdanu potvrdu kliničkog značaja i prognostičke efikasnosti analiziranih teksturalnih i morfometrijskih parametara se preporučuje dodatna provera rezultata kroz korišćenje proširene grupe pacijenata i eksternu validaciju.

Kao limitirajući faktor u ovoj studiji treba se uzeti u obzir i činjenica da pan-citokeratin koktel antitela AE1/AE3 boji i normalne i maligne epitelne ćelije. U cilju prevazilaženja ovog ograničenja, na osnovu morfološkog kriterijuma, izvršen je odabir segmenata slika na kojima su dominirale maligne ćelije. Shodno tome, ispitivanje raspodele intenziteta pan-citokeratinskog imunobojenja, u ovoj disertaciji, odgovaralo je analizi obrazaca rasta i stepena diferencijacije malignih ćelija. Ipak, treba imati u vidu da neophodni proces odabira reprezentativnih segmenata slike od strane stručnog patologa inkorporira subjektivnost u računarsku analizu koja je inače po svojoj prirodi objektivna.

7. ZAKLJUČAK

Na osnovu postavljenih hipoteza i rezultata analize distribucije intenziteta pan-citokeratinskog imunobojenja i prognostičkog učinka teksturalnih parametara na slikama odvojenih intenziteta sive skale histopatoloških uzoraka tumora dojke izvedeni su sledeći zaključci:

- 1) Podelom intenziteta piksela na sedam segmenata ovaj rad je po prvi put istraživao detaljnu prostornu raspodelu eksprimirane količine pan-citokeratina na epitelnim ćelijama karcinoma dojke. Većina piksela (59%) je imala vrlo nizak intenzitet (preko 220), 35% je imalo srednji intenzitet (130 - 220) dok je samo 6% piksela tamniji opseg intenziteta od 0 - 130. Pokazano je da intenzitet od 220 čini granicu između specifičnog i nespecifičnog pan-citokeratinskog bojenja pošto je analiza prostornog rasporeda intenziteta pan-citokeratinskog bojenja pokazala da tamniji pikseli sve do intenziteta 220 boje maligne ćelije, dok svetliji pikseli intenziteta preko 220 boje samo stromalni prostor koji se nalazi između klastera malignih ćelija.
- 2) Prognostički učinak teksturalnih GLCM i parametara statistike prvog reda nije bio ravnomerno raspoređen kod različitih intenziteta pan-citokeratinskog bojenja. Srednji i niski intenziteti pan-citokeratinskog imunobojenja u rasponu 120 - 200 su bili dominantni izvor prognostičkih informacija dok visoki intenziteti imunobojenja nisu sadržali prognostički značajne informacije. Ovim je potvrđen prognostički značaj nediferenciranih epitelnih ćelija koje najčešće eksprimiraju srednju i nisku količinu pan-citokeratina.
- 3) Podela histograma na podintervale i segmentacija slika u funkciji raspodele intenziteta bojenja je dovela do poboljšanja prognostičke pouzdanosti pri čemu je najizraženije poboljšanje zabeleženo za GLCM parametre IDM, ASM, ENT i COR čiji je učinak unapređen čak za dve kategorije mereno na AUC skali. Ovi teksturalni GLCM parametri su pokazali i najbolje prognostičke performanse sa ostvarenim vrlo dobrim učinkom u celokupnom intervalu intenziteta pri čemu su

prva dva indikatora povećane a ENT je indikator smanjene verovatnoće za pojavu udaljenih metastaza.

- 4) Prognošički učinak teksturalnih parametara ostvaren na slikama sive skale nadmašuje učinak demografskih i standardno korišćenih kliničko-patoloških prognošičkih parametara. Klinički značaj poboljšanja prognošičke pouzdanosti ishoda raka dojke se zasniva na potencijalnom omogućavanju prilagođavanja terapije svakom pojedinačnom pacijentu (personalizovana medicina) i rezultirajućem produženju preživljavanja pacijenta.

Na osnovu postavljenih hipoteza i rezultata analize morfometrijskih parametara na binarizovanim slikama histopatoloških uzoraka tumora dojke izvedeni su sledeći zaključci:

- 1) Utvrđeno je da klasteri malignih ćelija u tumoru dojke pokazuju ogromnu morfometrijsku raznolikost u parametrima veličine, oblika, broja i ekspresije citokeratina. Kombinacijom šest pragova binarizacije, pet pragova cirkularnosti i četiri praga veličine, dobijeno je 120 setova slika koji su omogućili detaljno ispitivanje raspodele prognošičkih informacija u odnosu na morfometrijske karakteristike malignih klastera.
- 2) Maligni klasteri visoke cirkularnosti su imali veći prognošički značaj od klastera niske cirkularnosti.
- 3) Maligni klasteri manjih dimenzija su imali veći prognošički značaj od klastera većih dimenzije.
- 4) Prognošički značaj je isključivo bio koncentrisan u malignim plažama sa manje citokeratina, dok plaže sa većom količinom eksprimiranih citokeratina nisu imale prognošički značaj. Ovim je opet potvrđen prognošički značaj slabije diferenciranih malignih ćelija.
- 5) Celokupna analiza prognošičkog značaja malignih klastera u odnosu na njihovu morfometriju je pokazala da su najveći prognošički značaj imale najmanje čestice najviše cirkularnosti koje su vizuelno mikroskopski identifikovane kao

pojedinačne maligne epitelne ćelije rasute po stromi tumora. Njihovo prisustvo je bilo indikator povećanog rizika razvoja metastaza a njihov broj je definisan kao najefikasniji i nezavisan prognostički parametar. Klinički značaj ovih ćelija povećava činjenica da se one mogu identifikovati, kvantifikovati i analizirati ne samo računarskom analizom već i vizuelnim mikroskopskim pregledom.

- 6) Prognostički učinak pomenutih morfometrijskih parametara epitelnih malignih plaža ostvaren na binarizovanim slikama histopatoloških uzoraka tumora dojke je nadmašio prognostički učinak standardno korišćenih demografskih i kliničko-patoloških prognostičkih parametara. Ovo ukazuje da nakon buduće dodatne validacije na dodatnim grupama pacijentkinja, postoji potencijal za kliničku primenu dobijenih rezultata u pružanju objektivne podrške u odlučivanju o izboru terapije kod pacijentkinja sa rakom dojke.

8. LITERATURA

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021 Feb 4;
2. Skrining Srbija - Publikacije [Internet]. [cited 2021 Apr 13]. Available from: <https://www.skriningsrbija.rs/srl/skrining-raka-dojke/publikacije/>
3. Global Cancer Observatory [Internet]. [cited 2022 Feb 14]. Available from: <https://gco.iarc.fr/>
4. Rajković N, Vujasinović T, Kanjer K, Milošević NT, Nikolić-Vukosavljević D, Radulovic M. Prognostic biomarker value of binary and grayscale breast tumor histopathology images. *Biomark Med.* 2016 Oct 1;10(10):1049–59.
5. Lønning PP, Knappskog S, Staalesen V, Chrisanthar R, Lillehaug JR. Breast cancer prognostication and prediction in the postgenomic era. Vol. 18, *Annals of Oncology.* Elsevier; 2007. p. 1293–306.
6. Cianfrocca M, Goldstein LJ. Prognostic and Predictive Factors in Early-Stage Breast Cancer. *Oncologist.* 2004 Nov 1;9(6):606–16.
7. Castellano G, Bonilha L, Li LM, Cendes F. Texture analysis of medical images. Vol. 59, *Clinical Radiology.* Elsevier; 2004. p. 1061–9.
8. He L, Long LR, Antani S, Thoma GR. Histology image analysis for carcinoma detection and grading. *Comput Methods Programs Biomed.* 2012 Sep;107(3):538–56.
9. Elston CW, Gresham GA, Rao GS, Žebro T, Haybittle JL, Houghton J, et al. The cancer research campaign (King's/Cambridge trial for early breast cancer: Clinico-pathological aspects. *Br J Cancer.* 1982;45(5):655–69.
10. Schnitt SJ. Classification and prognosis of invasive breast cancer: From morphology to molecular taxonomy. *Mod Pathol.* 2010 May 3;23(2):60–4.
11. Carlson B. Oncotype DX Test Offers Guidance For Women Debating Chemotherapy. *Biotechnol Healthc.* 2006 Aug;3(4):12–4.
12. Orucevic A, Chen J, McLoughlin JM, Heidel RE, Panella T, Bell J. Is the TNM Staging System for Breast Cancer Still Relevant in the Era of Biomarkers and Emerging Personalized Medicine for Breast Cancer - An Institution's 10-year

- Experience. *Breast J.* 2015 Mar 1;21(2):147-54.
13. Wiechmann L, Sampson M, Stempel M, Jacks LM, Patil SM, King T, et al. Presenting features of breast cancer differ by molecular subtype. *Ann Surg Oncol.* 2009 Oct 11;16(10):2705-10.
 14. Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J Natl Cancer Inst.* 2006 Sep 6;98(17):1183-92.
 15. Vranes V, Rajković N, Li X, Plataniotis KN, Raković NT, Milovanović J, et al. Size and Shape Filtering of Malignant Cell Clusters within Breast Tumors Identifies Scattered Individual Epithelial Cells as the Most Valuable Histomorphological Clue in the Prognosis of Distant Metastasis Risk. *Cancers (Basel).* 2019 Oct 22;11(10):1615.
 16. Dunn JM, Hveem T, Pretorius M, Oukrif D, Nielsen B, Albrechtsen F, et al. Comparison of nuclear texture analysis and image cytometric DNA analysis for the assessment of dysplasia in Barrett's oesophagus. *Br J Cancer.* 2011 Oct 11;105(8):1218-23.
 17. Laurinavicius A, Laurinaviciene A, Dasevicius D, Elie N, Benoît Plancoulaine B, Bor C, et al. Digital image analysis in pathology: Benefits and obligation. *Anal Cell Pathol.* 2012;35:75-8.
 18. Laurinavicius A, Plancoulaine B, Laurinaviciene A, Herlin P, Meskauskas R, Baltrusaityte I, et al. A methodology to ensure and improve accuracy of Ki67 labelling index estimation by automated digital image analysis in breast cancer tissue. *Breast Cancer Res.* 2014 Apr 6;16(2):3390.
 19. Angell HK, Gray N, Womack C, Pritchard DI, Wilkinson RW, Cumberbatch M. Digital pattern recognition-based image analysis quantifies immune infiltrates in distinct tissue regions of colorectal cancer and identifies a metastatic phenotype. *Br J Cancer.* 2013 Sep 17;109(6):1618-24.
 20. Pribic J, Vasiljevic J, Kanjer K, Konstantinovic ZN, Milosevic NT, Vukosavljevic DN, et al. Fractal dimension and lacunarity of tumor microscopic images as prognostic indicators of clinical outcome in early breast cancer. *Biomark Med.* 2015;9(12):1279-90.
 21. Aeffner F, Zarella M, Buchbinder N, Bui M, Goodman M, Hartman D, et al.

- Introduction to digital image analysis in whole-slide imaging: A white paper from the digital pathology association. *J Pathol Inform.* 2019 Jan 1;10(1):9.
22. Vujasinovic T, Pribic J, Kanjer K, Milosevic NT, Tomasevic Z, Milovanovic Z, et al. Gray-Level Co-Occurrence Matrix Texture Analysis of Breast Tumor Images in Prognosis of Distant Metastasis Risk. *Microsc Microanal.* 2015;21(3):646–54.
 23. Rajković N, Kolarević D, Kanjer K, Milošević NT, Nikolić-Vukosavljević D, Radulovic M. Comparison of Monofractal, Multifractal and gray level Co-occurrence matrix algorithms in analysis of Breast tumor microscopic images for prognosis of distant metastasis risk. *Biomed Microdevices.* 2016 Oct 1;18(5):1–13.
 24. Kolarević D, Vujasinović T, Kanjer K, Milovanović J, Todorović-raković N, Nikolić-Vukosavljević D, et al. Effects of different preprocessing algorithms on the prognostic value of breast tumour microscopic images. *J Microsc.* 2018;270(1):17–26.
 25. Boxberg M, Kuhn PH, Reiser M, Erb A, Steiger K, Pickhard A, et al. Tumor Budding and Cell Nest Size Are Highly Prognostic in Laryngeal and Hypopharyngeal Squamous Cell Carcinoma. *Am J Surg Pathol.* 2019 Mar 1;43(3):303–13.
 26. Hinck L, Näthke I. Changes in cell and tissue organization in cancer of the breast and colon. *Curr Opin Cell Biol.* 2014;26(1):87–95.
 27. Oldenhuis CNAM, Oosting SF, Gietema JA, de Vries EGE. Prognostic versus predictive value of biomarkers in oncology. *Eur J Cancer.* 2008 May;44(7):946–53.
 28. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: Correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science (80-).* 1987;235(4785):182–91.
 29. Hayes DF, Isaacs C, Stearns V. Prognostic factors in breast cancer: Current and new predictors of metastasis. Vol. 6, *Journal of Mammary Gland Biology and Neoplasia.* *J Mammary Gland Biol Neoplasia*; 2001. p. 375–92.
 30. Bardou VJ, Arpino G, Elledge RM, Osborne CK, Clark GM. Progesterone receptor status significantly improves outcome prediction over estrogen receptor status alone for adjuvant endocrine therapy in two large breast cancer databases. *J Clin Oncol.* 2003 May 15;21(10):1973–9.

31. Perez EA, Suman VJ, Davidson NE, Martino S, Kaufman PA, Lingle WL, et al. HER2 testing by local, central, and reference laboratories in the NCCTG N9831 Intergroup Adjuvant Trial. *J Clin Oncol*. 2004 Jul 15;22(14_suppl):567-567.
32. Shachar SS, Muss HB. Internet tools to enhance breast cancer care. *npj Breast Cancer*. 2016 Dec 14;2(1):1-4.
33. Turashvili G, Brogi E. Tumor Heterogeneity in Breast Cancer. *Front Med*. 2017;4(DEC):227.
34. Phung MT, Tin Tin S, Elwood JM. Prognostic models for breast cancer: A systematic review. Vol. 19, *BMC Cancer*. BioMed Central Ltd.; 2019. p. 1-18.
35. Albain KS, Allred DC, Clark GM. Breast cancer outcome and predictors of outcome: are there age differentials? *J Natl Cancer Inst Monogr*. 1994 Jan 1;(16):35-42.
36. Nixon AJ, Neuberg D, Hayes DF, Gelman R, Connolly JL, Schnitt S, et al. Relationship of patient age to pathologic features of the tumor and prognosis for patients with stage I or II breast cancer. *J Clin Oncol*. 1994;12(5):888-94.
37. Saez RA, McGuire WL, Clark GM. Prognostic factors in breast cancer. *Semin Surg Oncol*. 1989;5(2):102-10.
38. Carter CL, Allen C, Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. *Cancer*. 1989;63(1):181-7.
39. Fisher ER, Anderson S, Redmond C, Fisher B. Pathologic findings from the national surgical adjuvant breast project protocol B-06 10-year pathologic and clinical prognostic discriminants. *Cancer*. 1993;71(8):2507-14.
40. Fisher B, Bauer M, Wickerham DL, Redmond CK, Fisher ER, Cruz AB, et al. Relation of number of positive axillary nodes to the prognosis of patients with primary breast cancer. An NSABP update. *Cancer*. 1983;52(9):1551-7.
41. Rosen PP, Groshen S, Kinne DW, Norton L. Factors influencing prognosis in node-negative breast carcinoma: Analysis of 767 T1NoMo/T2NoMo patients with long-term follow-up. *J Clin Oncol*. 1993;11(11):2090-100.
42. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology*. 1991;19(5):403-10.
43. Bloom HJ, Richardson WW. Histological grading and prognosis in breast cancer

- a study of 1409 cases of which 359 have been followed for 15 years. *Br J Cancer*. 1957;11(3):359-77.
44. Contesso G, Mouriesse H, Friedman S, Genin J, Sarrazin D, Rouesse J. The importance of histologic grade in long-term prognosis of breast cancer: a study of 1,010 patients, uniformly treated at the Institut Gustave-Roussy. *J Clin Oncol*. 1987;5(9):1378-86.
 45. Harris LN, Ismaila N, McShane LM, Andre F, Collyar DE, Gonzales-Angulo AM, et al. Use of Biomarkers to Guide Decisions on Adjuvant Systemic Therapy for Women With Early-Stage Invasive Breast Cancer: American Society of Clinical Oncology Clinical Practice Guideline. *J Clin Oncol*. 2016 Apr 1;34(10):1134-50.
 46. Nassar A, Radhakrishnan A, Cabrero IA, Cotsonis GA, Cohen C. Intratumoral heterogeneity of immunohistochemical marker expression in breast carcinoma: A tissue microarray-based study. *Appl Immunohistochem Mol Morphol*. 2010 Oct;18(5):433-41.
 47. Harvey JM, Clark GM, Osborne CK, Allred DC. Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. *J Clin Oncol*. 1999;17(5):1474-81.
 48. Nadji M, Gomez-Fernandez C, Ganjei-Azar P, Morales AR. Immunohistochemistry of estrogen and progesterone receptors reconsidered: Experience with 5,993 breast cancers. *Am J Clin Pathol*. 2005;123(1):21-7.
 49. A Borg A. HER-2/neu amplification predicts poor survival in node positive carcinoma of the breast. *Cancer Res*. 1990;50:4332-7.
 50. Lobo J, See EYS, Biggs M, Pandit A. An insight into morphometric descriptors of cell shape that pertain to regenerative medicine. Vol. 10, *Journal of Tissue Engineering and Regenerative Medicine*. John Wiley and Sons Ltd; 2016. p. 539-53.
 51. Jara-Lazaro AR, Thamboo TP, Teh M, Tan PH. Digital pathology: Exploring its applications in diagnostic surgical pathology practice. Vol. 42, *Pathology*. Lippincott Williams and Wilkins; 2010. p. 512-8.
 52. Al-Janabi S, Huisman A, Van Diest PJ. Digital pathology: Current status and future perspectives. Vol. 61, *Histopathology*. *Histopathology*; 2012. p. 1-9.

53. Potts SJ, Young GD, Voelker FA. The role and impact of quantitative discovery pathology. Vol. 15, *Drug Discovery Today*. 2010. p. 943-50.
54. Meijer GA, Beliën JAM, Van Diest PJ, Baak JPA. Image analysis in clinical pathology. Vol. 50, *Journal of Clinical Pathology*. BMJ Publishing Group; 1997. p. 365-70.
55. Webster JD, Dunstan RW. Whole-Slide Imaging and Automated Image Analysis: Considerations and Opportunities in the Practice of Pathology. *Vet Pathol*. 2014 Jan 3;51(1):211-23.
56. Cornish TC, Swapp RE, Kaplan KJ. Whole-slide imaging: Routine pathologic diagnosis. Vol. 19, *Advances in Anatomic Pathology*. *Adv Anat Pathol*; 2012. p. 152-9.
57. Fónyad L, Krenács T, Nagy P, Zalatnai A, Csomor J, Sági Z, et al. Validation of diagnostic accuracy using digital slides in routine histopathology. *Diagn Pathol*. 2012 Mar 31;7(1).
58. Horbinski C, Fine JL, Medina-Flores R, Yagi Y, Wiley CA. Telepathology for intraoperative neuropathologic consultations at an Academic Medical Center: A 5-year report. *J Neuropathol Exp Neurol*. 2007 Aug;66(8):750-9.
59. Adamczyk A, Niemiec J, Ambicka A, Mucha-Małecka A, Ryš J, Mituš J, et al. Survival of breast cancer patients according to changes in expression of selected markers between primary tumor and lymph node metastases. *Biomark Med*. 2016 Mar 1;10(3):219-28.
60. Tseng SCG, Jarvinen MJ, Nelson WG, Huang JW, Woodcock-Mitchell J, Sun TT. Correlation of specific keratins with different types of epithelial differentiation: Monoclonal antibody studies. *Cell*. 1982 Sep 1;30(2):361-72.
61. Vranes V, Vujasinović T, Rajković N, Kanjer K, Milošević NT, Radulovic M. Analysis of Spatial Distribution and Prognostic Value of Different Pan Cytokeratin Immunostaining Intensities in Breast Tumor Tissue Sections. *Int J Mol Sci*. 2020 Jun 22;21(12):4434.
62. Rajković N, Li X, Plataniotis KN, Kanjer K, Radulovic M, Milošević NT. The Pan-Cytokeratin Staining Intensity and Fractal Computational Analysis of Breast Tumor Malignant Growth Patterns Prognosticate the Occurrence of Distant Metastasis. *Front Oncol*. 2018 Aug 30;8(AUG):348.

63. Di Cataldo S, Ficarra E. Mining textural knowledge in biological images: Applications, methods and trends. Vol. 15, Computational and Structural Biotechnology Journal. Elsevier B.V.; 2017. p. 56–67.
64. Tambasco M, Eliasziw M, Magliocco AM. Morphologic complexity of epithelial architecture for predicting invasive breast cancer survival. *J Transl Med.* 2010 Dec 31;8(1):140.
65. Todorović-Raković N, Jovanović D, Nešković-Konstantinović Z, Nikolić-Vukosavljević D. Prognostic value of HER2 gene amplification detected by chromogenic in situ hybridization (CISH) in metastatic breast cancer. *Exp Mol Pathol.* 2007 Jun 1;82(3):262–8.
66. Todorovic-Rakovic N, Jovanovic D, Neskovic-Konstantinovic Z, Nikolic-Vukosavljevic D. Comparison between immunohistochemistry and chromogenic in situ hybridization in assessing HER-2 status in breast cancer. *Pathol Int.* 2005 Jun 1;55(6):318–23.
67. Jacquemier J, Spyratos F, Esterni B, Mozziconacci MJ, Antoine M, Arnould L, et al. SISH/CISH or qPCR as alternative techniques to FISH for determination of HER2 amplification status on breast tumors core needle biopsies: A multicenter experience based on 840 cases. *BMC Cancer.* 2013 Jul 22;13(1):351.
68. Li X, Plataniotis KN. A Complete Color Normalization Approach to Histopathology Images Using Color Cues Computed From Saturation-Weighted Statistics. *IEEE Trans Biomed Eng.* 2015 Jul 1;62(7):1862–73.
69. Converting Image Formats - Learn ImageJ [Internet]. [cited 2021 Apr 24]. Available from: <https://sites.google.com/site/learnimagej/tutorials/converting-image-formats>
70. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Vol. 9, *Nature Methods.* Nat Methods; 2012. p. 671–5.
71. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: An open-source platform for biological-image analysis. Vol. 9, *Nature Methods.* Nature Publishing Group; 2012. p. 676–82.
72. Otsu N. Threshold Selection Method From Gray-Level Histograms. *IEEE Trans Syst Man Cybern.* 1979;SMC-9(1):62–6.
73. Sankur B. Survey over image thresholding techniques and quantitative

- performance evaluation. *J Electron Imaging*. 2004 Jan 1;13(1):146.
74. Lee SU, Yoon Chung S, Park RH. A comparative performance study of several global thresholding techniques for segmentation. *Comput Vision, Graph Image Process*. 1990 Nov 1;52(2):171-90.
 75. Chaubey AK. Comparison of The Local and Global Thresholding Methods in Image Segmentation. undefined. 2016;
 76. Ridler TW, Calvard S. Picture Thresholding Using An Iterative Selection Method. *IEEE Trans Syst Man Cybern*. 1978;SMC-8(8):630-2.
 77. Prabusankarlal KM, Thirumoorthy P, Manavalan R. Assessment of combined textural and morphological features for diagnosis of breast masses in ultrasound. 2011;
 78. Adur J, Carvalho HF, Cesar CL, Casco VH. Nonlinear optical microscopy signal processing strategies in cancer. Vol. 13, *Cancer Informatics*. Libertas Academica Ltd.; 2014. p. 67-76.
 79. Kim TY, Choi HJ, Hwang HG, Choi HK. Three-dimensional texture analysis of renal cell carcinoma cell nuclei for computerized automatic grading. *J Med Syst*. 2010 Aug;34(4):709-16.
 80. Aggarwal N, K. Agrawal R. First and Second Order Statistics Features for Classification of Magnetic Resonance Brain Images. *J Signal Inf Process*. 2012 May 30;03(02):146-53.
 81. Hartati S, Harjoko A. Selection Mammogram Texture Descriptors Based on Statistics Properties Backpropagation Structure. Vol. 11, *IJCSIS) International Journal of Computer Science and Information Security*. 2013.
 82. Julesz B. Experiments in the visual perception of texture. *Sci Am*. 1975;232(4):34-43.
 83. Haralick RM, Dinstein I, Shanmugam K. Textural Features for Image Classification. *IEEE Trans Syst Man Cybern*. 1973;SMC-3(6):610-21.
 84. Pantic I, Paunovic J, Perovic M, Cattani C, Pantic S, Suzic S, et al. Time-dependent reduction of structural complexity of the buccal epithelial cell nuclei after treatment with silver nanoparticles. *J Microsc*. 2013 Dec 1;252(3):286-94.
 85. Pantic I, Pantic S, Paunovic J, Perovic M. Nuclear entropy, angular second moment, variance and texture correlation of thymus cortical and medullar

- lymphocytes: Grey level co-occurrence matrix analysis. *An Acad Bras Cienc.* 2013 Sep;85(3):1063–72.
86. Vujasinovic T, Pribic J, Kanjer K, Milosevic NT, Tomasevic Z, Milovanovic Z, et al. Gray-Level Co-Occurrence Matrix Texture Analysis of Breast Tumor Images in Prognosis of Distant Metastasis Risk. *Microsc Microanal.* 2015 Apr 1;21(3):646–54.
 87. Larroza A, Bodí V, Moratal D. Texture Analysis in Magnetic Resonance Imaging: Review and Considerations for Future Applications. In: *Assessment of Cellular and Organ Function and Dysfunction using Direct and Derived MRI Methodologies.* InTech; 2016.
 88. Pantić I V. Univerzitet u Beogradu, Medicinski Fakultet. Fraktalna i teksturalna analiza strukturne kompleksnosti nukleusnog hromatina u postnatalnom razvoju i starenju, doktorska disertacija. 2013.
 89. Longstaff D, Longstaff D, Walker R, Walker RF, Jackway P, Jackway P. Improving Co-occurrence Matrix Feature Discrimination. *PROC DICTA'95, 3 RD Int Conf Digit IMAGE Comput Tech Appl.* 1995;643--648.
 90. Chen S, Zhao M, Wu G, Yao C, Zhang J. Recent advances in morphological cell image analysis. Vol. 2012, *Computational and Mathematical Methods in Medicine.* *Comput Math Methods Med;* 2012.
 91. Lepekhin EA, Walmod PS, Berezin A, Berezin V, Bock E. Evaluation of cell morphology. Vol. 161, *Methods in molecular biology (Clifton, N.J.). Methods Mol Biol;* 2001. p. 85–100.
 92. Metzler V, Lehmann T, Bienert H, Mottaghy K, Spitzer K. Scale-independent shape analysis for quantitative cytology using mathematical morphology. *Comput Biol Med.* 2000 May 15;30(3):135–51.
 93. Soll DR, Voss E, Varnum-Finney B, Wessels D. “Dynamic morphology system”: A method for quantitating changes in shape, pseudopod formation, and motion in normal and mutant amoebae of *Dictyostelium discoideum.* *J Cell Biochem.* 1988;37(2):177–92.
 94. Helmy IM, Abdel Azim AM. Efficacy of ImageJ in the assessment of apoptosis. *Diagn Pathol.* 2012 Feb 6;7(1):1–6.
 95. Jiang H, Zhu WJ, Li J, Chen QJ, Liang WB, Gu YQ. Quantitative histological

- analysis and ultrastructure of the aging human testis. *Int Urol Nephrol*. 2014 May 1;46(5):879–85.
96. Sołtys Z, Ziaja M, Pawliński R, Setkowicz Z, Janeczko K. Morphology of reactive microglia in the cerebral cortex. Fractal analysis and complementary quantitative methods. *J Neurosci Res*. 2001 Jan 1;63(1):90–7.
 97. Behnam-Motlagh P, Grankvist K, Henriksson R, Engström KG. Response in shape and size of individual p31 cancer cells to cisplatin and ouabain: A computerized image analysis of cell halo characteristics during continuous perfusion. *Cytometry*. 2000 Jul 1;40(3):198–208.
 98. Bray MAP, Adams WJ, Geisse NA, Feinberg AW, Sheehy SP, Parker KK. Nuclear morphology and deformation in engineered cardiac myocytes and tissues. *Biomaterials*. 2010 Jul 10;31(19):5143–50.
 99. Nafe R, Glienke W, Schlote W, Schneider B. EGFR gene amplification in glioblastomas: Is there a relationship with morphology of tumor cell nuclei and proliferative activity? *Anal Quant Cytol Histol*. 2001 Apr 1;23(2):135–43.
 100. Payne CM, Hicks MJ, Bjore CG, Kibler R. Ultrastructural morphometric analysis of nuclear contour irregularity in normal cord and adult blood: Correlation with distinct lymphocyte subpopulations. *Diagn Clin Immunol*. 1987 Jan 1;5(1):41–53.
 101. Soltys Z, Orzyłowska-Sliwinska O, Zaremba M, Orłowski D, Piechota M, Fiedorowicz A, et al. Quantitative morphological study of microglial cells in the ischemic rat brain using principal component analysis. *J Neurosci Methods*. 2005 Jul 15;146(1):50–60.
 102. Rocchi MBL, Sisti D, Albertini MC, Teodori L. Current trends in shape and texture analysis in neurology: Aspects of the morphological substrate of volume and wiring transmission. Vol. 55, *Brain Research Reviews*. 2007. p. 97–107.
 103. Smith TG, Lange GD, Marks WB. Fractal methods and results in cellular morphology - Dimensions, lacunarity and multifractals. Vol. 69, *Journal of Neuroscience Methods*. 1996. p. 123–36.
 104. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med*. 1999 Mar 16;130(6):515–24.
 105. Alemayehu D, Zou KH. Applications of ROC Analysis in Medical Research. Recent Developments and Future Directions. Vol. 19, *Academic Radiology*. Elsevier;

2012. p. 1457-64.
106. Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett.* 2006 Jun 1;27(8):861-74.
 107. Yerushalmy J. Statistical problems in assessing methods of medical diagnosis, with special reference to X-ray techniques. *Public Health Rep.* 1947 Oct 3;62(40):1432-49.
 108. Habibzadeh F, Habibzadeh P, Yadollahie M. On determining the most appropriate test cut-off value: The case of tests with continuous results. *Biochem Medica.* 2016 Oct 1;26(3):297-307.
 109. Kumar R, Indrayan A. Receiver operating characteristic (ROC) curve for medical researchers. Vol. 48, *Indian Pediatrics.* Springer; 2011. p. 277-87.
 110. Ma H, Bandos AI, Gur D. On the use of partial area under the ROC curve for comparison of two diagnostic tests. *Biometrical J.* 2015 Mar 1;57(2):304-20.
 111. Hoo ZH, Candlish J, Teare D. What is an ROC curve? *Emerg Med J.* 2017 Jun 1;34(6):357-9.
 112. Park SH, Goo JM, Jo CH. Receiver operating characteristic (ROC) curve: Practical review for radiologists. Vol. 5, *Korean Journal of Radiology.* Korean Radiological Society; 2004. p. 11-8.
 113. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival Analysis Part I: Basic concepts and first analyses. Vol. 89, *British Journal of Cancer.* Nature Publishing Group; 2003. p. 232-8.
 114. Kaplan EL, Meier P. Nonparametric Estimation from Incomplete Observations. *J Am Stat Assoc.* 1958;53(282):457-81.
 115. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard S V., et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. *Br J Cancer.* 1977;35(1):1-39.
 116. Stel VS, Dekker FW, Tripepi G, Zoccali C, Jager KJ. Survival analysis ii: Cox regression. *Nephron - Clin Pract.* 2011 Oct;119(3).
 117. Prentice RL. Introduction to Cox (1972) Regression Models and Life-Tables. In Springer, New York, NY; 1992. p. 519-26.
 118. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival Analysis Part III: Multivariate data analysis - Choosing a model and assessing its adequacy and fit.

- Vol. 89, British Journal of Cancer. Nature Publishing Group; 2003. p. 605–11.
119. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: A new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res.* 2004 Nov 1;10(21):7252–9.
 120. Efron B. Bootstrap Methods: Another Look at the Jackknife. *Ann Stat.* 1979 Jan;7(1):1–26.
 121. Brankovićmagić M V., Nikolićvukosavljević DB, Neškovićkonstantinović ZB, Kanjer KS, Spužić I V. Variations in the content of steroid receptors in breast cancer: Comparison between primary tumors and metastatic lesions. *Acta Oncol (Madr).* 1992;31(6):629–33.
 122. Prat A, Pineda E, Adamo B, Galván P, Fernández A, Gaba L, et al. Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast.* 2015 Nov 1;24:S26–35.
 123. Wang L-W, Qu A-P, Yuan J-P, Chen C, Sun S-R, Hu M-B, et al. Computer-Based Image Studies on Tumor Nests Mathematical Features of Breast Cancer and Their Clinical Prognostic Value. Brusic V, editor. *PLoS One.* 2013 Dec 12;8(12):e82314.
 124. Parry DA, Steinert P M. Intermediate filaments: molecular architecture, assembly, dynamics and polymorphism. *Q Rev Biophys.* 1999 May;32(2):99–187.
 125. Sterlekov SV, Herrmann H, Aebi U. Molecular architecture of intermediate filaments. *Bioessays.* 2003 Mar 1;25(3):243–51.
 126. Barak V, Goike H, Panaretakis K W, Einarsson R. Clinical utility of cytokeratins as tumor markers. *Clin Biochem.* 2004 Jul;37(7):529–40.
 127. Kirfel J, Magin T M, Reichelt J. Keratins: a structural scaffold with emerging functions. *Cell Mol Life Sci.* 2003 Jan 1;60(1):56–71.
 128. Osborn M, Weber K. Intermediate filaments: cell-type-specific markers in differentiation and pathology. *Cell.* 1982;31(2 Pt 1):303–6.
 129. Fuchs E, Weber K. Intermediate filaments: structure, dynamics, function, and disease. *Annu Rev Biochem.* 1994;63:345–82.
 130. Moll R, Frank WW, Schiller DL, Geiger B, Krepler R. The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells. *Cell.* 1982;31(1):11–24.
 131. Sredni ST, Neves JI, Camargo B de, Caballero OLD de, Soares FA. Pan-cytokeratin

- immunoexpression in Wilms' tumors: a simple approach for understanding tumor epithelial differentiation. *Sao Paulo Med J.* 2004 Jul 1;122(4):181-3.
132. Lane E, Alexander C. Use of keratin antibodies in tumor diagnosis. undefined. 1990;
 133. Chu PG, Weiss LM. Keratin expression in human tissues and neoplasms. *Histopathology.* 2002;40(5):403-39.
 134. Rajkovic N, Li X, Plataniotis KN, Kanjer K, Radulovic M, Milošević NT. The Pan-Cytokeratin staining intensity and fractal computational analysis of breast tumor malignant growth patterns prognosticate the occurrence of distant metastasis. *Front Oncol.* 2018;8(AUG).
 135. Tambasco M, Magliocco AM. Relationship between tumor grade and computed architectural complexity in breast cancer specimens. *Hum Pathol.* 2008 May 1;39(5):740-6.
 136. Brotherick I, Robson CN, Browell DA, Shenfine J, White MD, Cunliffe WJ, et al. Cytokeratin Expression in Breast Cancer: Phenotypic Changes Associated With Disease Progression. *Cytometry.* 1998;32:301-8.
 137. Xu Q, Yuan JP, Chen YY, Zhang HY, Wang LW, Xiong B. Prognostic Significance of the Tumor-Stromal Ratio in Invasive Breast Cancer and a Proposal of a New Ts-TNM Staging System. *J Oncol.* 2020;2020.
 138. Gujam FJ, Edwards J, Mohammed ZMA, Going JJ, McMillan DC. The relationship between the tumour stroma percentage, clinicopathological characteristics and outcome in patients with operable ductal breast cancer. *Br J Cancer.* 2014;111(1):157-65.
 139. Hansen TF, Kjæer-Frifeldt S, Lindebjerg J, Rafaelson SR, Jensen LH, Jakobsen A, et al. Tumor-stroma ratio predicts recurrence in patients with colon cancer treated with neoadjuvant chemotherapy. *Acta Oncol.* 2018 Apr 3;57(4):528-33.
 140. Fridman WH, Galon J, Pages F, Tartour E, Sautes-Fridman C, Kroemaer G. Prognostic and predictive impact of intra- and peritumoral immune infiltrates. *Cancer Res.* 2011 Sep 1;71(17):5601-5.
 141. Park JH, Richards CH, McMillan DC, Horgan PG, Roxburgh CSD. The relationship between tumour stroma percentage, the tumour microenvironment and survival in patients with primary operable colorectal cancer. *Ann Oncol Off J Eur Soc*

- Med Oncol. 2014 Mar;25(3):644-51.
142. Rye IH, Trinh A, Saetersdal AB, Nebdal D, Lingjaerde OC, Almendro V, et al. Intratumor heterogeneity defines treatment-resistant HER2+ breast tumors. *Mol Oncol*. 2018 Nov 1;12(11):1838-55.
 143. Beca F, Polyak K. Intratumor Heterogeneity in Breast Cancer. *Adv Exp Med Biol*. 2016;882:169-89.
 144. Gerdes MJ, Gökmen-Polar Y, Sui Y, Pang AS, LaPlante N, Harris AL, et al. Single-cell heterogeneity in ductal carcinoma in situ of breast. *Mod Pathol* 2018 313. 2017 Nov 17;31(3):406-17.
 145. Liang F, Cao W, Wang Y, Li L, Zhang G, Wang Z. The prognostic value of tumor budding in invasive breast cancer. *Pathol Res Pract*. 2013 May 1;209(5):269-75.
 146. Zhou Z, Lu ZR. Molecular imaging of the tumor microenvironment. *Adv Drug Deliv Rev*. 2017 Apr 1;113:24-48.
 147. Freeman MR, Li Q, Chung LWK. Can Stroma Reaction Predict Cancer Lethality? *Clin Cancer Res*. 2013 Sep 15;19(18):4905-7.
 148. Pinto G, Alhaiek AAM, Amadi S, Qattan AT, Crawford M, Radulovic M, et al. Systematic Nucleo-Cytoplasmic Trafficking of Proteins Following Exposure of MCF7 Breast Cancer Cells to Estradiol. *J Proteome Res*. 2014 Feb 7;13(2):1112-27.
 149. Rizki A, Bissell MJ. Homeostasis in the breast: It takes a village. *Cancer Cell*. 2004 Jul 1;6(1):1-2.
 150. Beck AH, Sangoi AR, Leung S, Marinelli RJ, Nielsen TO, Van De Vijver MJ, et al. Imaging: Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. *Sci Transl Med*. 2011;3(108):1-11.
 151. Wang LW, Qu AP, Yuan JP, Chen C, Sun SR, Hu MB, et al. Computer-based image studies on tumor nests mathematical features of breast cancer and their clinical prognostic value. *PLoS One*. 2013 Dec 12;8(12).
 152. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. *J Clin Invest*. 2009 Jun 1;119(6):1420-8.
 153. Voutsadakis IA. Prognostic role of tumor budding in breast cancer. *World J Exp Med*. 2018 Sep 7;8(2):12.
 154. Grigore AD, Jolly MK, Jia D, Farach-Carson MC, Levine H. Tumor Budding: The

- Name is EMT. Partial EMT. *J Clin Med* 2016, Vol 5, Page 51. 2016 Apr 29;5(5):51.
155. Gujam FJA, McMillan DC, Mohammed ZMA, Edwards J, Going JJ. The relationship between tumour budding, the tumour microenvironment and survival in patients with invasive ductal breast cancer. *Br J Cancer* 2015 1137. 2015 Aug 11;113(7):1066-74.
 156. Roxanis I, Colling R, Kartsonaki C, Green AR, Rakha EA. The significance of tumour microarchitectural features in breast cancer prognosis: a digital image analysis. *Breast Cancer Res* 2018 201. 2018 Feb 5;20(1):1-11.
 157. Fehm T, Sagalowsky A, Clifford E, Beitsch P, Saboorian H, Euhus D, et al. Cytogenetic Evidence That Circulating Epithelial Cells in Patients with Carcinoma Are Malignant. *Clin Cancer Res*. 2002;8(7).
 158. Allard WJ, Matera J, Miller MC, Repollet M, Connelly MC, Rao C, et al. Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases. *Clin Cancer Res*. 2004 Oct 15;10(20):6897-904.
 159. Smith J, Mathisen AF, Funch Richardt N, Vander Plaetsen AS, Van Nieuwerburgh F, Stender H, et al. Feasibility of single-cell analysis of model cancer and foetal cells in blood after isolation by cell picking. *Tumor Biol*. 2019 Feb 1;41(2).
 160. Tan W, Liang G, Xie X, Jiang W, Tan L, Sanders AJ, et al. Incorporating MicroRNA into Molecular Phenotypes of Circulating Tumor Cells Enhances the Prognostic Accuracy for Patients with Metastatic Breast Cancer. *Oncologist*. 2019 Nov 1;24(11):e1044-54.
 161. Alvarez Cubero MJ, Lorente JA, Robles-Fernandez I, Rodriguez-Martinez A, Puche JL, Serrano MJ. Circulating Tumor Cells: Markers and Methodologies for Enrichment and Detection. *Methods Mol Biol*. 2017;1634:283-303.
 162. Kent P, Boyle E, Cancelliere C, Cassidy JD, Kongsted A, Kongsted A. A conceptual framework for prognostic research. *BMC Med Res Methodol*. 2020 Jun 29;20(1):172.
 163. Riley R, Windt D van der, Croft P, Moons K. Prognosis research in healthcare: concepts, methods, and impact. 2019.
 164. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin*

Epidemiol. 1996 Dec;49(12):1373-9.

165. Steyerberg EW, Harrell FE, Borsboom GJJM, Eijkemans MJC, Vergouwe Y, Habbema JDF. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol.* 2001;54(8):774-81.

Biografija autora

Veličko Vraneš rođen je 04.08.1967. godine u Bijeljini. Osnovnu školu "Sedam sekretara SKOJ-a" i srednju elektrotehničku školu "Nikola Tesla" (smer elektrotehničar za medicinske uređaje) je završio u Beogradu. Diplomirao je na Fizičkom Fakultetu u Beogradu 2000. godine na smeru Fizičar za opštu fiziku, sa prosečnom ocenom 8,40.

Nakon završenih studija seli se u Dominikansku Republiku gde radi kao profesor fizike. Na privatnom univerzitetu UNIBE (Universidad Iberoamericana), u periodu od 2002 do 2006 predaje predmete Fizika I, II i III studentima inženjerstva, kao i predmet Biofizika, studentima medicine. Od 2000. godine je zaposlen i na privatnom univerzitetu INTEC (Instituto Tecnológico de Santo Domingo) gde predaje predmete: Diferencijalni i Integralni račun, Fizika I, II i III, Osnovi Fizike I i II i Biofizika studentima medicine i biotehnologije.

U toku 2004 godine je odbranio magistarski rad iz oblasti pedagogije visokog obrazovanja na univerzitetu UNIBE. Interdisciplinarne doktorske studije iz oblasti biofizike upisuje na Beogradskom univerzitetu 2015. godine i započinje saradnju sa Medicinskim Fakultetom u Beogradu i Institutom za onkologiju i radiologiju Srbije u oblasti računarske analize medicinskih slika.

Učestvovao je u radu različitih naučnih skupova: "8th Regional Biophysics Conference" 2018. godine u Sloveniji, međunarodna konferencija "Arithmetic Methods In Mathematical Physics and Biology" 2018. godine u Poljskoj i "22nd International Conference on Control Systems and Computer Science" 2019. godine u Rumuniji.

Autor je više radova objavljenim u međunarodnim časopisima i na međunarodnim skupovima, kao i laboratorijskih praktikuma korišćenih u praktičnoj nastavi na univerzitetima UNIBE i INTEC. Autor je udžbenika "Fundamentos de Biofísica" koji se koristi kao primarni udžbenik u nastavi biofizike za studente medicine i biotehnologije na univerzitetu INTEC.

Od 2018. godine je član Društva biofizičara Srbije i udruženja SODOFI (Sociedad Dominicana de Física). Govori srpski, španski i engleski jezik.

Изјава о ауторству

Име и презиме аутора Величко Вранеш

Број индекса 117/2015

Изјављујем

да је докторска дисертација под насловом

ОПТИМИЗАЦИЈА РАЧУНАРСКЕ АНАЛИЗЕ МОРФОЛОГИЈЕ ЕПТЕЛНИХ МАЛИГНИХ ПЛАЖА
ТУМОРА ДОЈКЕ У ЦИЉУ ПОБОЉШАЊА ПРОГНОЗЕ ИСХОДА БОЛЕСТИ

- резултат сопственог истраживачког рада;
- да дисертација у целини ни у деловима није била предложена за стицање друге дипломе према студијским програмима других високошколских установа;
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио/ла интелектуалну својину других лица.

Потпис аутора

У Београду, 15.05.2022.



Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора Величко Вранеш

Број индекса 117/2015

Студијски програм Биофизика

Наслов рада ОПТИМИЗАЦИЈА РАЧУНАРСКЕ АНАЛИЗЕ МОРФОЛОГИЈЕ ЕПИТЕЛНИХ МАЛИГНИХ ПЛАЖА ТУМОРА ДОЈКЕ У ЦИЉУ ПОБОЉШАЊА ПРОГНОЗЕ ИСХОДА БОЛЕСТИ

Ментор Проф. др Небојша Милошевић

Коментор Научни саветник др Марко Радловић

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла ради похрањена у **Дигиталном репозиторијуму Универзитета у Београду**.

Дозвољавам да се објаве моји лични подаци везани за добијање академског назива доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

Потпис аутора

У Београду, 15.05.2022.



Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

ОПТИМИЗАЦИЈА РАЧУНАРСКЕ АНАЛИЗЕ МОРФОЛОГИЈЕ ЕПИТЕЛНИХ МАЛИГНИХ ПЛАЖА ТУМОРА ДОЈКЕ У ЦИЉУ ПОБОЉШАЊА ПРОГНОЗЕ ИСХОДА БОЛЕСТИ

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигиталном репозиторијуму Универзитета у Београду и доступну у отвореном приступу могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство (CC BY)
2. Ауторство – некомерцијално (CC BY-NC)
3. Ауторство – некомерцијално – без прерада (CC BY-NC-ND)
4. Ауторство – некомерцијално – делити под истим условима (CC BY-NC-SA)
5. Ауторство – без прерада (CC BY-ND)
6. Ауторство – делити под истим условима (CC BY-SA)

(Молимо да заокружите само једну од шест понуђених лиценци.
Кратак опис лиценци је саставни део ове изјаве).

Потпис аутора

У Београду, 15.05.2022.



1. **Ауторство.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце, чак и у комерцијалне сврхе. Ово је најслободнија од свих лиценци.

2. **Ауторство – некомерцијално.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела.

3. **Ауторство – некомерцијално – без прерада.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела. У односу на све остале лиценце, овом лиценцом се ограничава највећи обим права коришћења дела.

4. **Ауторство – некомерцијално – делити под истим условима.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца не дозвољава комерцијалну употребу дела и прерада.

5. **Ауторство – без прерада.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца дозвољава комерцијалну употребу дела.

6. **Ауторство – делити под истим условима.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца дозвољава комерцијалну употребу дела и прерада. Слична је софтверским лиценцама, односно лиценцама отвореног кода.