

RAPORT ȘTIINȚIFIC

1. Date proiect

Contract nr. 46PCE/08.01.2025

Perioada raportare: etapa nr. 1/2025

Titlu proiect: Aplicarea delta-radiomicii utilizând inteligența artificială pentru personalizarea tratamentului în oncologia ORL

Website proiect: https://cfiz.uoradea.ro/?page_id=1473

2. Descrierea științifică cu punerea în evidență a rezultatelor etapei anuale și gradul de realizare a obiectivelor

2.1. Obiectivul proiectului și planul de realizare

Obiectivul principal al proiectului constă în utilizarea analizei radiomice și a datelor holomice pentru a prezice rezultatul radioterapiei, riscul de recurență precum și timpul probabil de apariție a recurenței în pacienții cu cancere din sfera ORL. Astfel, caracteristicile radiomice îmbogățite cu informații clinice vor fi utilizate ca input pentru rețelele de învățare profundă instruite și validate pe date retrospective de la o cohortă mare de pacienți ORL tratați la clinicile Amethyst din Cluj-Napoca. Pentru atingerea obiectivului vor fi identificate caracteristici imagistice tumorale detaliate și robuste (radiomica) folosind învățarea profundă, suplimentate cu date non-imagistice pentru caracterizarea complexă a pacientului (holomica) (figura 1).

Conform planului de realizare a proiectului, activitățile aferente primului an de implementare (Etapa I – an 2025) respectiv rezultatele verificabile ale activității au fost următoarele:

✓ **Activități:**

- *Identificarea parametrilor radiomici si holomici (clinici) relevanți în cancerele din sfera ORL.*
- *Selecția datelor pacienților. Stabilirea bazei de date pentru modelul radiomic.*

✓ **Rezultate verificabile ale activității:**

- *Identificarea parametrilor relevanți în cancerele din sfera ORL.*
- *Baza de date utilizată pentru antrenarea algoritmului AI.*
- *Publicarea rezultatelor în reviste științifice indexate în Web of Science.*

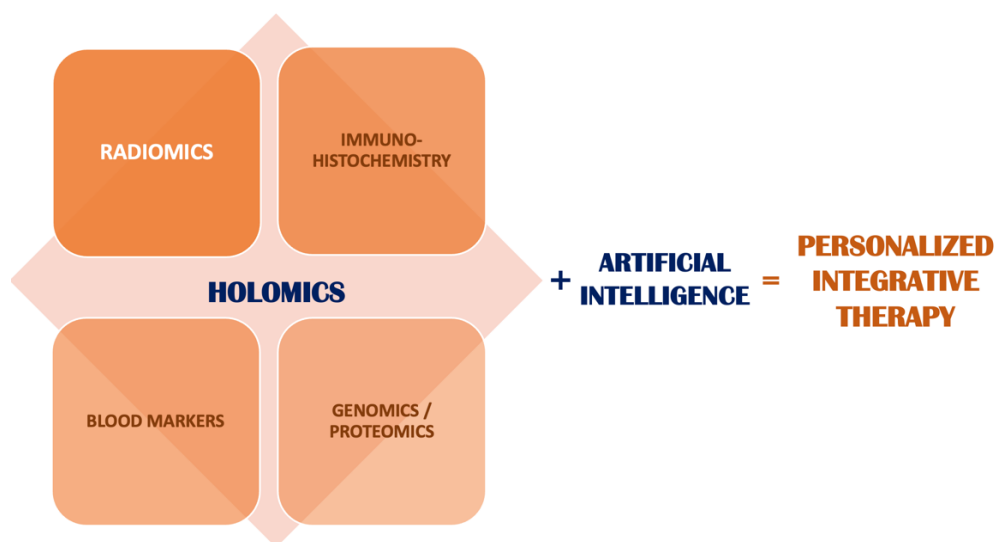


Figura 1. Ilustrarea schematică a ecuației terapiei personalizate integrative în oncologie

Tabelul 1. Planul de realizare a proiectului pentru etapa 1 (anul 2025)

Denumire activitate	Rezultate estimative verificabile ale activității	Grad îndeplinire
<i>Act. 1.1 - Recenzia literaturii</i>	<i>Cercetarea literaturii pentru identificarea parametrilor radiomici specifici cancerelor ORL</i>	Îndeplinit
<i>Act. 1.2 - Identificare corelații parametri</i>	<i>Cercetarea literaturii pentru determinarea corelațiilor între parametrii radiomici identificați anterior și parametrii radiobiologici specifici tumorilor ORL</i>	Îndeplinit
<i>Act. 1.3 - Identificare date clinice</i>	<i>Identificarea datelor clinice relevante care influențează rezultatul tratamentului</i>	Îndeplinit
<i>Act. 1.4 - Extragerea datelor clinice</i>	<i>Extragerea datelor pacienților din baza de date clinica - imagini CT și informații clinice specifice</i>	Îndeplinit
<i>Act. 1.5 - Identificare pacienți</i>	<i>Identificarea grupului de pacienți care necesită imagistica suplimentară pentru evaluarea sfârșitului de tratament</i>	Îndeplinit
<i>Act. 1.6 - Anonimizarea datelor</i>	<i>Anonimizarea datelor în vederea protecției datelor personale a pacienților</i>	Îndeplinit
<i>Act. 1.7 - Normalizarea datelor</i>	<i>Normalizarea imaginilor pentru a asigura uniformitatea extragerii parametrilor</i>	Îndeplinit

2.2. Evaluarea aplicațiilor de machine learning în radioterapia cancerelor din sfera ORL. Identificarea parametrilor radiomici relevanți.

Un prim pas în realizarea obiectivelor a constat în recenzia literaturii de specialitate, în vederea identificării aplicațiilor de inteligență artificială (IA) (machine learning - ML) în imagistica și radioterapia aferente cancerelor din sfera ORL. S-a realizat astfel o evaluare a bazelor de date Medline/PubMed și Web of Science pentru a identifica studii clinice care raportează utilizarea IA / ML în cancerele ORL pentru oricare dintre următoarele aspecte clinice: detectare / clasificare, segmentare a imaginii și predicție a răspunsului la tratament / distribuției dozei. Dintre cele 2395 de studii identificate, 42 au îndeplinit criteriile de eligibilitate (figura 2).

O evaluare critică prin analiza riscului de bias a tuturor studiilor, utilizând instrumentul Cochrane de evaluare a riscului de bias, a fost efectuată pentru a evalua calitatea analizei. Deși pentru majoritatea domeniilor riscul de bias este scăzut, au fost identificate unele provocări legate de compararea studiilor eligibile dintr-o anumită secțiune, din cauza numărului diferit de pacienți evaluați, a diferitelor modele de învățare automată comparate, a diferiților parametri extrași/raportați.

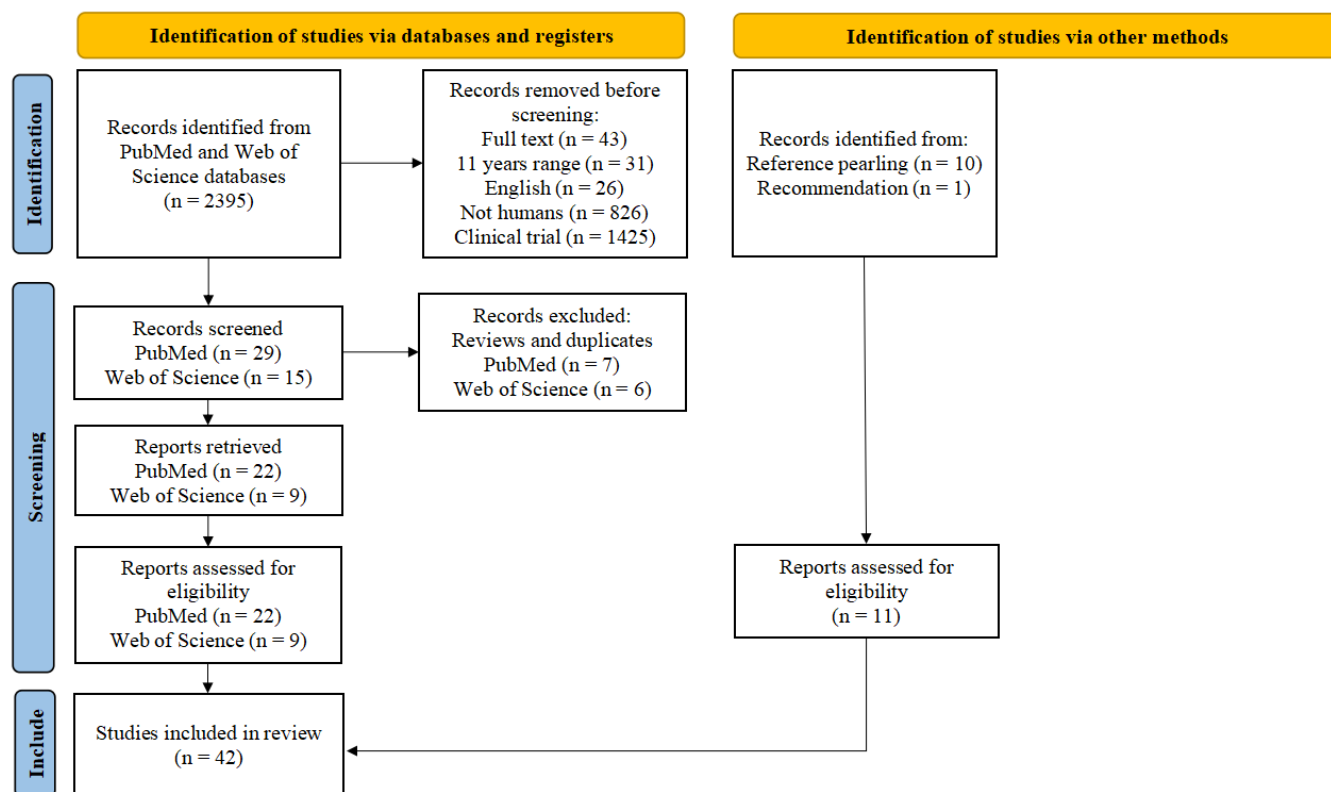


Figura 2. Diagrama PRISMA de recenzie sistematică a literaturii în vederea identificării studiilor care au analizat rolul ML în studii / trialuri clinice pentru tratamentul cancerelor din sfera ORL.

În plus, deoarece scopul analizei a fost de a demonstra rolul inteligenței artificiale în studiile clinice privind cancerele ORL, a fost analizată și completată o listă de verificare pentru inteligența artificială în imagistica medicală (CLAIM) pentru toate studiile eligibile din analiză pentru a indica transparența, nivelul de reproductibilitate și relevanța clinică a fiecărui studiu în

parte (figura 3). Scorul CLAIM mediu a fost de 32 (interval: 17-39; 44 de itemi verificați), arătând că majoritatea studiilor au raportat criteriile de includere/excludere, pre-procesarea datelor, subseturile de date, descrierea detaliată a metodelor utilizate pentru obținerea standardului de referință, proporțiile datelor din fiecare partiție (grup de antrenament/testare), descrierea modelelor ML, abordarea de învățare automată, indicatorii pentru performanța modelului ML și rezultatele acestora, ceea ce a dus la o calitate ridicată a raportării și a indicat relevanța clinică a rezultatelor acestora.

Studiile analizate au arătat că există un interes evident pentru creșterea nivelului de automatizare implicat în procesele de detectare, clasificare, segmentare a imaginilor și predicție a succesului tratamentului în tumorile ORL. Dispunerea de instrumente care pot ajuta profesioniștii din domeniul medical în aceste domenii permite reducerea variabilității inter- și intra-observator, ducând la un rezultat medical mai consistent.

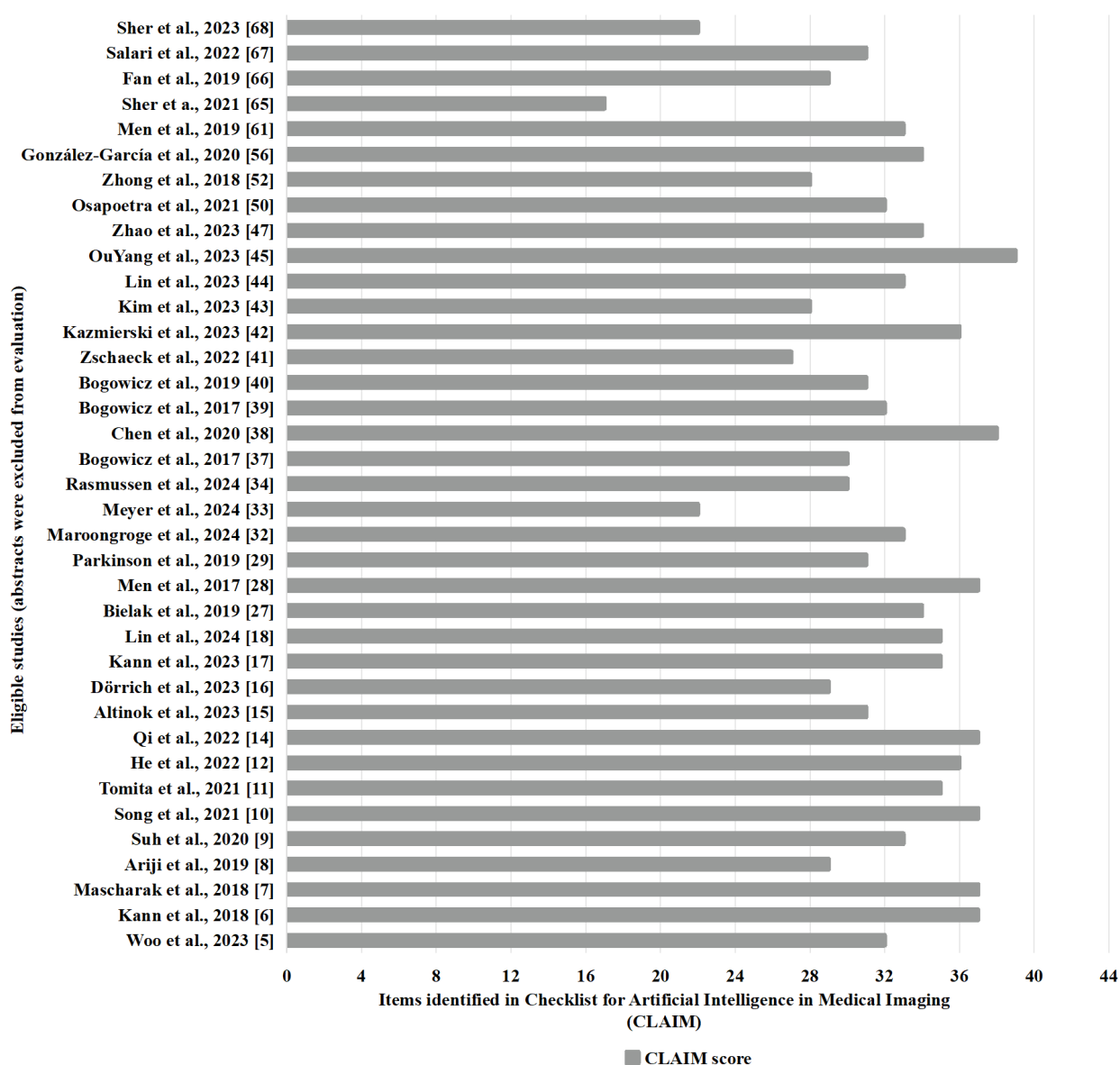


Figura 3. Scorul CLAIM (Artificial Intelligence in Medical Imaging) pentru studiile analizate.

Studiile care au investigat comparativ performanța diferiților algoritmi au concluzionat că modelele de deep learning care utilizează rețele neuronale cu straturi multiple sunt mai robuste și depășesc metodele de machine learning în dezvoltarea modelelor predictive. Alegerea algoritmului utilizat pentru procesarea datelor depinde de aplicația dezvoltată, deoarece în multe cazuri algoritmi mai tradiționali, precum KNN, SVM și regresia logistică, au produs rezultate comparabile cu algoritmi de deep learning. Selectarea metodei care oferă cel mai bun echilibru între acuratețea rezultatului și resursele de implementare ar putea face diferența dintre un studiu științific și o soluție implementabilă clinic.

De obicei, există o discrepanță mare între numărul de caracteristici radiomice luate în considerare și numărul de caracteristici utilizate în studiu. Unul dintre motive este că practic orice calcul bazat pe pixelii imaginii poate fi considerat o caracteristică radiomică, dar nu toate aceste calcule sunt capabile să ofere informații noi și utile pentru a îmbogăți setul de informații. Există, de asemenea, o tendință naturală a cercetătorilor de a favoriza caracteristicile care au o semnificație inteligibilă, care ar putea fi observate de profesioniști, dar care ar necesita mult timp pentru a le cuantifica valoarea. După cum este prezentat în tabelul 2, cele mai frecvent utilizate caracteristici radiomice au fost legate de textura imaginii (matricea de co-ocurență a nivelului de gri - GLCM, matricea lungimii de rulare a nivelului de gri - GLRLM, matricea lungimii zonei de nivel de gri - GLZLM, matricea de diferență a nivelului de gri de vecinătate - NGLDM, matricea de dependență a nivelului de gri - GLDM), indiferent de domeniul de aplicare al modelului ML.

Tabel 2. Cele mai utilizate caracteristici radiomice raportate în literatura de specialitate, clasificate în funcție de aplicațiile ML în managementul cancerelor ORL.

Obiectivul studiului	Semnătura radiomică (numărul de caracteristici radiomice)
Detectarea / clasificarea cancerelor ORL	
Identificarea statusului HPV folosind imagistica IRM (<i>Suh et al., 2020</i>)	$N_{\text{total}} = 1618$ $N_{\text{extract}} = 7$ GLCM_dist_1: autocorrelation, homogeneity GLCM_dist_2: entropy, correlation. GLCM_dist_3: entropy, correlation.
Diferențierea între adenomul pleomorf parotidian și adenolinfomul parotidian folosind secvențe IRM (T1WI, T2WI, T1-2WI) (<i>Song et al., 2021</i>)	$N_{\text{total}} = 429$ T1WI features $N_{\text{total}} = 441$ T2WI features Intensity histogram GLCM GLRLM Shape
Identificarea metastazelor ganglionilor limfatici cervicali de nivel I, II și I/II (<i>Tomita et al., 2021</i>)	$N_{\text{extract}} = 25$ Histogram features: skewness, kurtosis, entropy, energy. GLCM: homogeneity, energy, contrast, correlation, entropy, dissimilarity.

	GLRLM: short run emphasis, low and high gray-level run emphasis, short run low and high gray-level emphasis, run percentage. GLZLM: short-zone emphasis, low and high gray-level zone emphasis, short-zone low and high gray-level emphasis, zone length non-uniformity, Zone percentage. NGLDM: coarseness, contrast.
Clasificarea tumorilor parotide în 4 categorii utilizând imagini IRM (<i>He et al., 2022</i>)	$N_{\text{total}} = 3396$ $N_{\text{extract}} = 28$ Shape- and size-based features: flatness, elongation (using CE-T1WI and T2WI). First-order features: maximum, kurtosis, 90 percentile, uniformity, root mean square GLCM: entropy, contrast, inverse variance (using CE-T1WI and T2WI), maximum probability, Id, correlation, Idmm, Imc2 (using T2WI, T1WI, CE-T1WI) GLRLM: run entropy, gray level non-uniformity normalized GLSZM: high gray level zone emphasis (using CE-T1WI and T2WI), high gray level zone. GLDM: wavelet transformation, the Laplacian of Gaussian transformation
Pentru a diferenția între tumorile scuamoase de grad scăzut și ridicat (<i>Lin et al., 2022</i>)	$N_{\text{extract}} = 9$
Predictia statusului HPV in cancerile de orofaringe folosind imagini CT cu contrast (<i>Altinok et al., 2023</i>)	$N_{\text{total}} = 851$ $N_{\text{extract}} = 25$ Shape features: 2 most significant radiomic features (Sphericity and Max2DDiameterRow) Intensity features Texture: GLCM, GLRLM, GLSZM
Evaluarea controlului tumoral local	
Imagistica funcțională PET-CT pre-tratament pentru controlul tumoral local (<i>Bogowicz et al., 2017a</i>)	$N_{\text{total}} = 569$ Shape features Intensity features Texture: GLCM, NGTDM, GLSZM Wavelet transformation
Imagistica CT pre-tratament pentru controlul tumoral local (<i>Bogowicz et al., 2017b</i>)	$N_{\text{total}} = 317$ $N_{\text{extract}} = 9$ LLL sum entropy, LHL imc1, HHH large zone high gray-level emphasis, HHL contrast NGTDM, LLH difference variance, HLH complexity, Sum entropy, LHH mcc, LHL low gray-level emphasis

Predicția controlului loco-regional folosind imagini CT (<i>Bogowicz et al., 2019</i>)	$N_{\text{total}} = 285$ $N_{\text{extract}} = 7$ Tumor GLSZM zone entropy, Tumor LLL NGTDM complexity, Tumor LLL GLCM entropy, LN SD thickness LN spherical disproportion, LN major axis, LN histogram kurtosis
Valoarea prognostică după rezecția chirurgicală utilizând imagini CT cu substanță de contrast (<i>Chen et al., 2020</i>)	$N_{\text{extract}} = 36$ First-order features: histogram and geometry (skewness, kurtosis, entropy, sphericity, compacity) Second-order features: GLCM (homogeneity, entropy, energy, contrast, correlation, dissimilarity), NGLDM (contrast, coarseness) Third-order features: GLRLM, GLZLM
Predicția supraviețuirii globale la 2 ani, riscul de deces utilizând imagini CT (<i>Kazmierski et al., 2023</i>)	$N_{\text{total}} = 1316$ Shape features First-order features Texture features
Valoarea prognostică a supraviețuirii globale utilizând imaginile CT (<i>Lin et al., 2023</i>)	Shape features Histogram-based features Texture: GLCM, GLDM, GLRLM.
Predicția metastazelor nodulilor limfatici prin imagini CT (<i>Zhao et al., 2023</i>)	$N_{\text{total}} = 960$ Shape features First-order features Texture Image filter (Laplacian of Gaussian and wavelet transformation)
Răspunsul tumorii la tratament	
Predicția răspunsului tumoral la radioterapia radicală folosind radiomica de ecografie cantitativă (<i>Osapoetra et al., 2021</i>)	$N_{\text{total}} = 105$ GLCM (contrast, correlation, energy, homogeneity) 5 spectral features 20 Second-order features 80 Third-order features
Alte aplicații ale algoritmilor de machine learning in radioterapie	
Evaluarea eficienței de substituție a ENI (elective node irradiation) cu iradierea nodulilor limfatici suspicioși (<i>Sher et al., 2023</i>)	$N_{\text{total}} = 257$ Intensity features: minimum, maximum, mean, standard deviation, sum, median, skewness, kurtosis, variance Geometry features: volume, major diameter, minor diameter, eccentricity, elongation orientation, bounding box volume, perimeter Texture features: GLCM, GLRL - energy, entropy, correlation, contrast, texture variance, sum-mean, inertia, cluster shade, cluster prominence, homogeneity, max-probability, inverse variance

Abrevieri: N_{total} = număr total de caracteristici radiomice raportate, N_{extract} = număr total de caracteristici radiomice extrase.

Notă: Terminologia consacrată descrierii caracteristicilor radiomice a fost păstrată în limba engleză pentru a nu se pierde din înțeles / interpretare.

Rezultatele / concluziile cercetării:

- ✓ Majoritatea studiilor au confirmat utilitatea ML în aplicațiile clinice ale cancerelor ORL prin parametrii raportați: precizie, aria de sub curbă, specificitate, sensibilitate, coeficientul Dice.
- ✓ Modelele de deep learning care utilizează rețele neuronale cu straturi multiple au fost adesea alegera preferată, deoarece depășesc performanțele învățării automate în dezvoltarea modelelor predictive.
- ✓ Cu toate acestea, algoritmi tradiționali (KNN k-nearest neighbor, SVM support vector machine, regresia logistică) au arătat adesea rezultate comparabile cu cele ale învățării automate.
- ✓ Cele mai frecvent utilizate caracteristici radiomice au fost legate de textura imaginii, indiferent de aplicația modelului.
- ✓ Rezultatele studiilor clinice arată un rol de susținere al ML în gestionarea cancerelor ORL, fie că este vorba de diagnostic/stadializare sau tratament. Cele mai performante aplicații sunt acelea care îmbunătățesc algoritmi de învățare profundă aplicații imagistice cu date clinice pentru o abordare mai cuprinzătoare și mai bine integrată a tratamentului personalizat.

2.3. Identificarea parametrilor holomici / clinici. Evaluarea radiomică a parametrilor caracteristici cancerelor din sfera ORL.

După cum a fost menționat mai sus, pe lângă analiza radiomică a imaginilor anatomice, un factor esențial în evaluarea și prezicerea rezultatului tratamentului respectiv a probabilității de recidivă, îl reprezintă caracteristicile specifice tumorilor ORL, cum ar fi abilitatea de proliferare. Una dintre activitățile prevăzute pentru anul 2025, ca parte din obiectivele proiectului, a fost *cercetarea literaturii pentru determinarea corelațiilor între parametri radiomici identificați anterior și parametri radiobiologici specifici tumorilor ORL* (tabelul 1). În acest scop, am efectuat un studiu din care să reiasă posibilele corelații între parametri radiomici și cei aferenți proliferării tumorale în cancerele din sfera ORL.

În urma analizei literaturii de specialitate, ca parametru / biomarker reprezentativ pentru proliferare a fost ales nivelul de expresie a proteinei nucleare Ki-67, proteină care este prezentă în celulele active, fiind identificată drept unul dintre cei mai preciși biomarkeri în ceea ce privește prezicerea rezultatelor tratamentului pe baza ratei de proliferare ajutând, implicit, în stratificarea pacienților. Au fost astfel analizate cele mai recente cercetări privind studiile de învățare automată care au identificat semnături / caracteristici radiomice referitoare la proliferare, care să fie utilizate pentru prezicerea nivelurilor de expresie Ki-67 în tumorile ORL (figura 4).

Pe baza analizei studiilor identificate prin diagrama PRISMA, următoarele rezultate și observații sunt relevante proiectului de față:

- Dezvoltările continue ale tehnologiei imagistice și utilizarea învățării automate în procesarea imaginilor au permis stabilirea unor corelații între caracteristicile biologice și particularitățile imagistice în majoritatea tipurilor de cancer.

- Imagistica prin rezonanță magnetică a evoluat în ultimele decenii și a devenit o parte integrantă a evaluării anatomiei capului și gâtului, alăturându-se CT și PET/CT-ului, deja consacrate, pentru diagnostic, stadializare și monitorizarea răspunsului la tratament.
- Mai multe studii au evaluat potențialul tehnicilor imagistice menționate mai sus pentru a prezice comportamentul biologic al celulelor canceroase din sfera ORL, unele concentrându-se în mod specific pe aspecte legate de proliferarea tumorală prin prezicerea nivelurilor de expresie a Ki-67.

O caracteristică bazată pe imagini frecvent întâlnită în studiile IRM ponderat prin difuzie a fost corelația inversă dintre valorile de difuzie și expresia Ki-67, deoarece coeficienții aparenti de difuzie (ADC) mai mici au fost asociați cu niveluri ridicate de Ki-67, sugerând o celularitate ridicată. Deși anumiți parametri cantitativi (cum ar fi ADC) au arătat asocieri cu nivelurile de Ki-67, este probabil ca aceste măsurători să nu reprezinte optim tumora și să nu ilustreze pe deplin o imagine eterogenă.

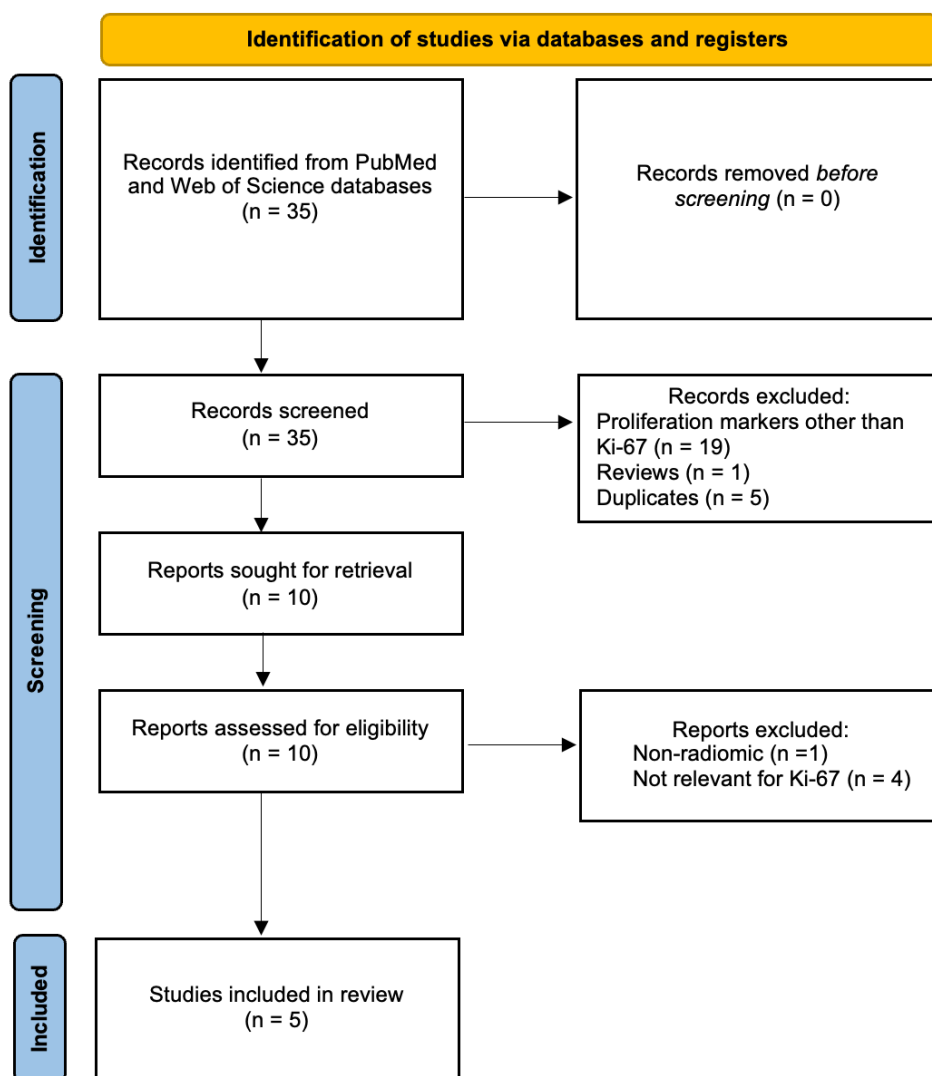


Figura 4. Diagrama PRISMA de recenzie sistematică a literaturii în vederea utilizării ML pentru stratificarea pacienților ORL în funcție de proliferarea tumorală indicată de biomarkerul Ki-67.

Pentru o analiză mai aprofundată a imaginilor diagnostice, au fost efectuate studii pentru a crea algoritmi de învățare automată care să cuprindă mai bine eterogenitatea tumorii. În timp ce majoritatea studiilor radiomice dezvoltate pentru predicția Ki-67 s-au bazat pe imagini IRM, unele și-au construit modelul radiomic pe imagini CT cu substanță de contrast (tabelul 3). În studiul lor, Zheng și colab. au antrenat și validat nomograma radiomică folosind două grupuri (expresie Ki-67 ridicată vs. scăzută, cu o valoare cut-off de 50%), obținând o performanță bună de discriminare cu aria de sub curbă (AUC) de 0,832 în setul de testare. Având în vedere natura retrospectivă a studiului lor, selecția pacienților (diferite subtipuri de cancere ORL) ar putea reprezenta un posibil bias, având în vedere variația valorii limită Ki-67 între diferite locații anatomice, așa cum arată studiile de imunohistochimie. Astfel, cohorte mai mari, cu subtipuri ORL bine reprezentate, ar oferi mai multă precizie asupra valorii prognostice a radiomicii bazate pe CT utilizând markeri de proliferare.

Pe lângă caracteristicile radiomice, factorii clinici s-au dovedit a fi completări valoroase ale modelelor ML pentru a le crește precizia. Aceste modele de fuziune au încorporat gradul histologic, infiltrațiile nodale, stadiul clinic T, starea ganglionilor limfatici raportată prin IRM, oferind astfel o abordare holomică mai complexă. Studiul radiomic realizat pe 265 de pacienți cu carcinom scuamos sinusal a identificat scorul radiomic, stadiul T, gradul histologic și Ki-67 ca predictorii independenți ai recurenței precoce, utilizând date IRM multiparametrice preoperatorii (Lin et al. 2023). În modelul radiomic, pacienții au fost stratificați în funcție de risc, utilizând software-ul X-tile pentru a găsi valorile limită ale nomogramei. Supraviețuirea fără recurență după 2 ani a fost estimată la 70,60% la pacienții cu risc scăzut și 6,30% la pacienții cu risc crescut ($p < 0,001$) în setul de antrenament, cu rezultate similare obținute în setul de validare: 69,40% vs 8,80% ($p < 0,001$). Frecvențele de recurență precoce observate s-au dovedit a fi în concordanță cu cele prezise de nomogramele radiomice, confirmând beneficiul clinic al modelului.

Cu un scop similar studiului menționat anterior, Wang și colab. au construit o nomogramă radiomică axată pe cancerele nazofaringiene pentru a evalua valoarea predictivă a indicelui Ki-67. Studiul lor, desfășurat în două centre, a inclus 133 de pacienți, datele dintr-un centru formând setul de antrenament, în timp ce al doilea centru a furnizat setul de testare. Pacienții au fost stratificați în grupuri cu risc scăzut vs. risc ridicat, pe baza unei valori limită de 50% pentru indicele KI-67. Analiza curbei de decizie a indicat o performanță foarte bună a nomogramei radiomice, cu o valoare predictivă bună a Ki-67 (AUC = 0,841 pentru setul de testare), demonstrând utilitatea clinică. Studiul a arătat, de asemenea, o corelație puternică între expresia Ki-67 și metastazele ganglionilor limfatici cervicali. Sub rezerva validării într-o cohortă mai mare, autorii recomandă utilizarea modelului lor ca instrument de diagnostic pentru planificarea tratamentului la pacienții cu cancer nazofaringian.

În studiul realizat de Chen și colab., pentru a ilustra ponderea caracteristicilor radiomice adăugate la datele clinice, o analiză a curbei ROC a arătat o creștere semnificativă a AUC după încorporarea caracteristicilor radiomice (cohorta de antrenament de la 0,737 la 0,916, $p < 0,001$; cohorta de validare internă de la 0,715 la 0,903, $p < 0,011$; cohorta de validare externă de la 0,654 la 0,885, $p < 0,001$). Aceste rezultate arată avantajul modelelor de fuziune față de cele clinice, ilustrând valoarea caracteristicilor radiomice în creșterea eficacității diagnostice și prezicerea robustă a nivelurilor de expresie Ki-67 la pacienții ORL cu carcinom cu celule scuamoase. În plus, studiul a confirmat asocierea puternică dintre stadiul T / starea ganglionilor limfatici și nivelurile de expresie Ki-67, demonstrată anterior de mai multe studii clinice.

Un studiu radiomic recent bazat pe imagini IRM, realizat pe 124 de pacienți cu carcinoame scuamoase orale, a analizat valoarea predictivă a Ki-67 pentru întreaga cohortă, precum și într-un

subgrup de 104 pacienți cu carcinoame orale cu celule scuamoase, indicând o utilitate clinică ridicată bazată pe adaptarea la datele clinice. Conform AUC atât în setul de antrenament (0,943; 0,960), cât și în cel de validare (0,801; 0,817) pentru întreaga cohortă, respectiv pentru cancerule orale, Ki-67 a fost identificat ca un excelent prezicător al prognosticului, confirmând utilizarea modelelor de nomogramă radiomică pentru predicția preoperatorie neinvazivă a indicelui Ki-67 la pacienții cu cancer ORL.

O evaluare indirectă a rolului Ki-67 în caracterizarea carcinoamelor ORL prin intermediul caracteristicilor radiomice a fost raportată într-un studiu care analizează asocierea dintre liganzii PD-1 tumoral și caracteristicile texturale pe baza imaginilor PET/CT. Corelațiile pozitive dintre expresiile PD-L1 și Ki-67 în studiile de imunohistochimie au justificat investigații suplimentare utilizând caracteristici texturale. Pe baza analizei curbei ROC, s-a demonstrat că Ki-67 este predictiv pentru expresii PD-L1 de $> 5\%$ (utilizat ca prag), concluzionând că nivelurile de Ki-67 pot oferi informații suplimentare la evaluarea nivelurilor de proteină a ligandului 1 al apoptozei în cancerule ORL.

O analiză a riscului de bias a studiilor eligibile incluse în această analiză, prin intermediul instrumentului Cochrane Risk of Bias, a fost efectuată pentru a evalua calitatea rezultatelor, concluzionând că pentru majoritatea domeniilor riscul de bias este scăzut, existând unele provocări cu privire la selecția algoritmului ML și validarea (performanța) modelului.

Tabelul 3. Studii radiomice în cancerul ORL pentru a prezice nivelurile de expresie Ki-67

Scopul studiului / nr. pacienți [ref]	Modul de determinare a Ki-67	Tehnica imagistică și caracteristicile radiomice	Corelații dintre caracteristicile radiomice și expresia Ki-67
Predicție a expresiei Ki-67 bazată pe nomogramă radiomică; 217 pacienți cu cancer ORL (140 pentru antrenare rețea; 77 pentru validare externă) (Zheng et al. 2023)	Imunohistochimie Valoare cut-off indice Ki-67: 50%	CT cu substanță de contrast. Regresia LASSO s-a utilizat pentru selecția parametrilor cheie ale modelului radiomic. 20 caracteristici selectate pentru Rad-score din 697 (17 <i>wavelet transformed features</i> , 2 <i>shape-based features</i> , 1 <i>first-order statistics feature</i>).	Nomograma radiomică (Rad-score + parametri clinici) a prezentat o valoare predictivă mai mare pentru Ki-67 decât modelul clinic. Analiza ROC a atribuit o valoare clinică mai mare modelului radiomic.
Predicția recurenței precoce în cancerul ORL bazată pe învățare profundă; 265 pacienți (165 pentru antrenare; 70 validare) (Lin et al. 2023)	Imunohistochimie (informație extrasă din datele pacienților)	IRM multiparametric. Regresia LASSO s-a aplicat pentru selecția parametrilor și construirea Rad-score. S-au extras 15 caracteristici radiomice din initialul de 2264 (<i>first-order features</i> , <i>texture and shape features</i>).	Rad-score, stadiul T, gradul și Ki-67 au fost predictorii independenți. Ki-67 a fost asociat cu prognosticul în modelul nomogramei radiomice.

Predictia indicelui Ki-67 în cancerul nazofaringian bazată pe nomogramă radiomică; 133 pacienți (105 pentru antrenare; 28 validare) (Wang et al. 2024)	Imunohistochimie Valoare cut-off indice Ki-67: 50%	IRM multiparametric. Regresia LASSO s-a aplicat pentru selecția parametrilor și construirea Rad-score. S-au extras 5 caracteristici radiomice din initialul de 1526 (<i>wavelet-filtered features</i>).	Nomograma a depășit modelul clinic ca performanță, cu o valoare predictivă bună pentru expresia Ki-67, așa cum este demonstrat de AUC = 0.841 pentru setul de validare.
Predictia preoperatorie bazată pe ML a expresiei Ki-67; 351 pacienți ORL (196 pentru antrenare; 84 validare internă; 71 validare externă) (Chen et al. 2024)	Imunohistochimie Valoare cut-off indice Ki-67: 50%	IRM multiparametric. SVM a fost selectat drept modelul ML optim (Rad-score). S-au extras 13 caracteristici radiomice din initialul de 1688 (<i>9 wavelet transformed features, 3 first-order statistical features and 1 filter transformed feature</i>).	Modelul de fuziune (Rad-score + date clinice) a diferențiat cu precizie între indicii Ki-67 în funcție de valoarea cut-off. Acuratețe: 85.2% (antrenare); 85.7% (validare internă); 84.5% (validare externă).
Predictia bazată pe radiomică a statusului Ki-67 în cancerul ORL; 124 pacienți (87 pentru antrenare; 37 validare) O subgrupă de 104 pacienți cu cancer oral (72 antrenare; 32 test) (Yang et al. 2025)	Imunohistochimie Valoare cut-off indice Ki-67: 50%	DCE-IRM. Regresia LASSO s-a aplicat pentru selecția parametrilor și construirea Rad-score. S-au extras 10 caracteristici radiomice pentru modelul predictiv în cohorta ORL; 14 caracteristici selectate pentru analiza subgrupeii cu cancer oral.	Radiomica DCE-IRM a prezentat rezultate excelente în prezicerea statusului Ki-67. Antrenare: AUC = 0.943 (ORL); 0.960 (cancer oral) Test: AUC = 0.801 (ORL); 0.817 (cancer oral).
Abrevieri: Rad-score = scor radiomic; LASSO = the least absolute shrinkage and selection operator; SVM = support vector machine; DCE = dynamic contrast-enhanced; ROC = receiver operating characteristics; AUC = area under the curve.			

Rezultatele / concluziile cercetării:

- ✓ Modelele bazate pe ML/deep learning pot prezice nivelurile de expresie Ki-67 în cancerul ORL și sunt capabile să asiste în luarea deciziilor clinice.
- ✓ În timp ce unele studii au identificat SVM (support vector machine) ca fiind algoritmul ML optim datorită echilibrului său dintre capacitatea de învățare și complexitate, altele au preferat regresia LASSO datorită superiorității sale în screening-ul caracteristicilor
- ✓ O caracteristică comună a studiilor actuale este numărul relativ mic de pacienți utilizați în învățarea/testarea acestor modele radiomice, dar cu rezultate bune care justifică o

- validare suplimentară în cohorte mai mari pentru a crește acuratețea și, implicit, valoarea lor predictivă.
- ✓ Deși majoritatea acestor studii radiomice sunt publicate recent și sunt limitate ca număr, rezultatele susțin rolul învățării automate în predicțiile rezultatelor clinice bazate pe rata de proliferare și justifică investigații suplimentare în domeniu.

2.4. Extragerea datelor clinice

Odată cu analiza literaturii de specialitate, au fost identificați și factorii clinici (date clinice) care influențează rezultatul tratamentului. Astfel, datele pacienților înrolați în studiul din cadrul proiectului conțin atât informații imagistice (din care se extrag caracteristicile radiomice) cât și date clinice personale (vârsta, sexul) dar și date relevante proprietăților tumorale individuale (stadializare), respectiv ale tratamentului cu radiație (doza prescrisă, regimul de fracționare, răspuns tumoral, toxicitate).

Toate aceste informații sunt incluse într-o bază de date care este în continuă creștere, datorită incluziunii unui număr tot mai mare de pacienți ORL, conform planului inițial. De asemenea, cohortele sunt împărțite în funcție de locația anatomică a tumorii (subgrupe tumorale) din sfera ORL: orofaringe, rinofaringe, nazofaringe, etc. Această decizie a fost luată în timpul colectării datelor, pentru a identifica posibilele diferențe radiomice între subgrupe și de a oferi o personalizare mai adecvată în privința regimului de radiație.

Datele imagistice sunt stocate în mediul clinic într-un format descris de standardul DICOM (Digital Imaging and Communications in Medicine). Imaginile formate conform acestui standard sunt trecute într-un sistem de arhivare de tip PACS (Picture Archiving and Communication System), acesta fiind o bază de date ce asigură siguranța datelor și accesul facil la informație doar pentru utilizatorii autentificați. Acest sistem este folosit în colaborare cu un sistem de tip RIS (Radiology Information System) care gestionează fluxul de lucru în radiologie și stochează datele textuale despre studiile imagistice. Informații complete despre pacienți care leagă datele imagistice de istoricul clinic mai larg, rezultatele de laborator și rezultatele tratamentelor sunt conținute în dosarul pacientului care este menținut într-un sistem de tip EHR (Electronic Health Record).

Pentru extragerea datelor clinice necesare, primul pas a fost identificarea grupului țintă de pacienți. Prin colaborare cu echipa clinică s-au definit criteriile de includere/excludere. Aceste criterii au ținut cont de localizarea anatomică a tumorii, de tratamentul aplicat, de rezultatele tratamentului și de numărul de investigații imagistice disponibile pentru pacient.

După definirea criteriilor de includere s-a rulat interogarea inițială pe sistemele RIS și EHR pentru a genera o listă de identificatori de pacienți și studii ce corespund criteriilor stabilite. Având lista de pacienți s-a trecut la preluarea imaginilor din sistemul de arhivare PACS. Comunicarea cu arhiva PACS a fost realizată folosind protocoalele DICOM C-FIND pentru a localiza studiile și C-MOVE pentru a transfera fișierele DICOM.

Antetele fișierelor de tip DICOM conțin etichete (tags) de metadate extinse care stochează informații despre pacienți, cum ar fi:

- (0010,0010) Numele pacientului
- (0010,0020) ID-ul pacientului
- (0010,0030) Data nașterii pacientului
- (0010,1010) Vârsta pacientului
- (0008,0050) Numărul de acces (Accession Number)

- (0008,0090) Numele medicului trimițător
- (0008,0080) Numele instituției
- (0008,0020) Data studiului.

În conformitate cu cerințele GDPR (General Data Protection Regulation), s-a efectuat procesul de anonimizare a datelor pentru a asigura securitatea informațiilor private ale pacienților. Numărul de ID al pacientului stocat în antetul fișierului imagistic a fost înlocuit cu un număr de identificare aparținând studiului de cercetare. Acest lucru a fost necesar pentru a putea face corelații între diferitele investigații imagistice pentru același pacient. Tabelul de corespondență între ID-ul pacientului din sistemul clinic și cel din studiul de cercetare a fost salvat și stocat de către cadrele medicale din clinica unde s-au efectuat investigațiile medicale și tratamentul. Restul de etichete cu date ce ar fi permis identificarea individuală a pacienților au fost șterse din antetele fișierelor DICOM. S-a efectuat o inspecție vizuală a imaginilor medicale pentru a confirma lipsa oricăror informații identificatoare a pacienților.

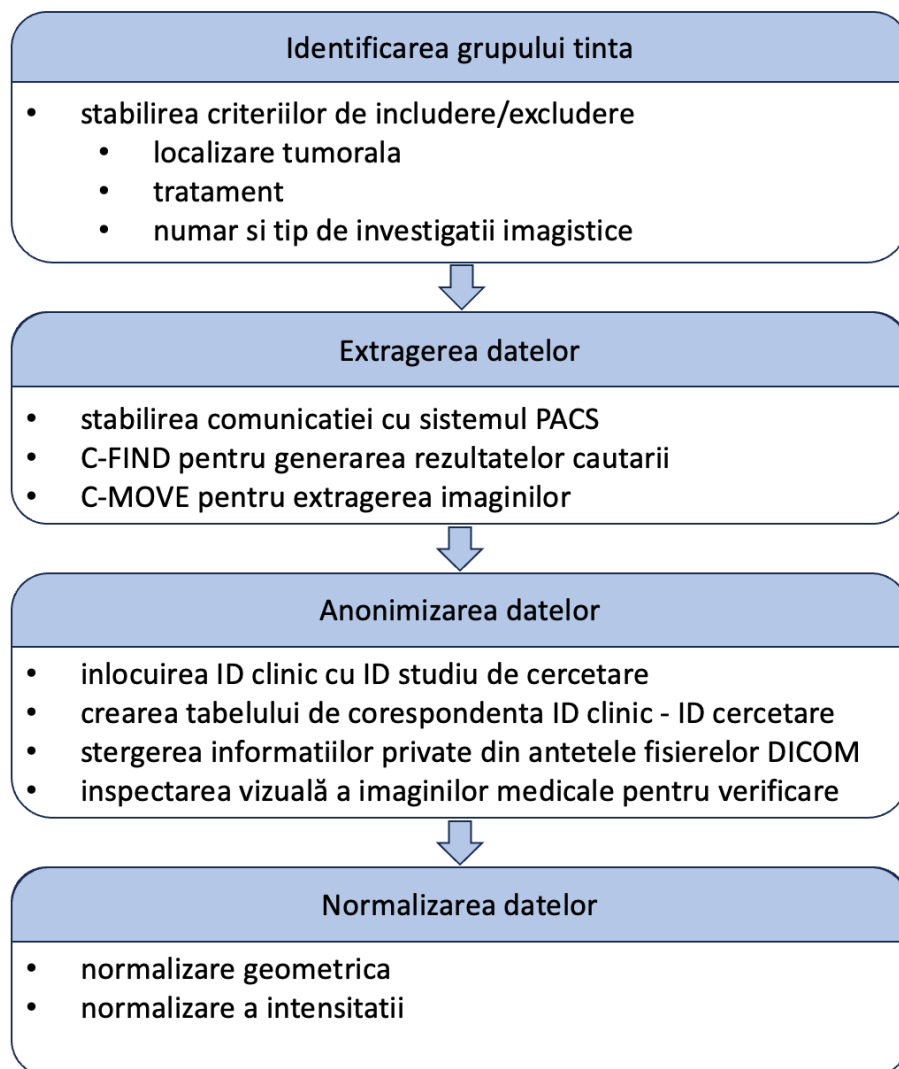


Figura 5. Pașii următori pentru identificarea datelor clinice ale grupului țintă

După obținerea setului de date anonimizat s-au inspectat datele din punct de vedere tehnic pentru a identifica diferențe tehnice datorate modalității de imagistică. Pentru a putea efectua un studiu transversal e nevoie ca toate datele imagistice să respecte aceleași standarde, astfel încât la măsurarea diferențelor între imagini să se obțină o diferență biologică reală și nu o diferență determinată de achiziționare imagistică diferită.

Pentru normalizarea geometrică a imaginilor ce au necesitat acest procedeu s-a executat reeșantionarea volumului folosind metoda de interpolare B-spline. Această metodă asigură o interpolare fină cu contururi relativ bine definite. Pentru normalizarea intensității s-a discutat cu cadrele medicale în vederea stabilirii nivelului optim de intensitate de tip HU (Hounsfield Unit) pentru a evidenția țesuturile de interes în imagine.

Pașii următori pentru identificarea datelor clinice ale grupului de pacienți cu cancer din sfera ORL înrolați în acest studiu sunt reprezentați în figura 5.

Având în vedere cele de mai sus, toate obiectivele au fost realizate conform planului inițial. Pe lângă obiectivele stabilite în propunerea de proiect, au fost inițiate o serie de activități adiționale, care vin în completarea ideilor inițiale (așa cum se arată în secțiunea 3. Sumar al progresului).

3. Sumar al progresului (livrabile realizate, indicatori de rezultat, diseminarea rezultatelor, justificare diferențe, dacă e cazul)

Activitățile propuse în cadrul proiectului au progresat conform planului inițial; livrabilele, indicatorii de rezultat și modul de diseminare a rezultatelor sunt specificați în Tabelul 4.

Tabel 4. Livrabile, indicatori de rezultat și diseminarea rezultatelor.

Livrabile realizate	Indicatori de rezultat	Diseminarea rezultatelor
Realizarea website-ului proiectului	Website proiect	Website proiect: https://cfiz.uoradea.ro/?page_id=1473
Analiza literaturii științifice cu privire la rolul inteligenței artificiale în imagistica medicală	Articol științific indexat WoS	L. Marcu, D. Marcu. <i>Examining the role of AI in cancer imaging through the lens of clinical studies</i> . Health & Technology (2025) DOI: https://doi.org/10.1007/s12553-025-01019-w Q3 (FI); AIS = 0,507 (Q4)
Analiza literaturii științifice cu privire la utilizarea inteligenței artificiale / machine learning în radioterapia cancerelor din sfera ORL / identificarea caracteristicilor radiomice relevante proiectului.	Articol științific indexat WoS	I.C. Costin, L. Marcu, D. Marcu, R. Zahu, O. Straciuc. <i>Is the use of machine learning in head and neck cancer radiotherapy supported by clinical trials?</i> Physica Medica 136:105036 (2025) DOI: 10.1016/j.ejmp.2025.105036 Q2 (FI); AIS = 0,715 (Q2)

Identificarea datelor clinice relevante care influențează rezultatul tratamentului.		
Analiza literaturii științifice cu privire la corelațiile dintre semnătura / caracteristicile radiomice ale cancerelor ORL respectiv parametrii radiobiologici specifici tumorilor ORL. Identificarea / confirmarea datelor clinice relevante care influențează rezultatul tratamentului.	Articol științific indexat WoS	L. Marcu, D. Marcu, I.C. Costin, R. Zahu, O. Straciuc. <i>From immunohistochemistry to machine learning-based patient stratification by tumour proliferation characteristics in head and neck cancer</i> . Critical Reviews in Oncology/Hematology 216:104987 (2025) DOI: 10.1016/j.critrevonc.2025.104987 Q1 (FI); AIS = 1,483 (Q1)
Evaluarea diferențelor de sex în răspunsul la tratament al pacienților cu cancer ORL și importanța personalizării terapiei.	Prezentare invitată conferința internațională de fizică medicală	L. Marcu. <i>Adding gender to therapy: the emerging need for gender-based studies for treatment optimisation in oncology</i> . World Congress on Medical Physics & Biomedical Engineering, Adelaide, 27 Sept-4 Oct (2025)
Extragerea datelor pacienților din baza de date clinica - imagini CT si informații clinice specifice. Identificarea grupului de pacienți care necesita imagistica suplimentara pentru evaluarea sfârșitului de tratament.	Realizare baza de date cu pacienți înrolați în studiul proiectului	În pregătire – diseminarea rezultatelor acestor activități se preconizează pentru anul 2 al proiectului (2026)
Anonimizarea datelor în vederea protecției datelor personale a pacienților		
Normalizarea imaginilor pentru a asigura uniformitatea extragerii parametrilor		
Prelucrarea semnalelor și modelarea de parametri în diagnosticul asistat în vederea utilizării informațiilor în tratamentul personalizat.	Teză de doctorat finalizată și susținută	D. Marcu – <i>Prelucrarea semnalelor și modelarea unor parametri în diagnosticul asistat și terapiile personalizate</i> – Universitatea din Oradea (2025)
Fluidizarea workflow-ului radioterapiei prin utilizarea AI în segmentarea imaginilor pacienților cu cancer ORL în vederea realizării planului de tratament.	Articol științific în lucru	Articol în lucru: I.C. Costin, L. Marcu, D. Marcu. <i>AI-based segmentation in head and neck cancer for personalisation and outcome optimisation in radiotherapy</i> .

4. Rezumat executiv al activităților realizate în perioada de implementare

Proiectul vizează cercetarea în domeniul oncologiei ORL în vederea personalizării tratamentului prin utilizarea inteligenței artificiale (machine learning) pentru procesarea imaginilor radiologice de planificare respectiv monitorizare a răspunsului la tratament. Astfel, pentru primul an al proiectului, s-a urmărit evaluarea statusului actual al cercetării internaționale în domeniu, atât din perspective ale aplicației inteligenței artificiale în imagistica medicală, cât și în radioterapia respectiv radiobiologia cancerelor din sfera ORL. De asemenea, s-a inițiat extragerea datelor pacienților din cohorta de studiu, anonimizarea datelor, normalizarea imaginilor și crearea bazei de date incluzând subgrupe de pacienți în funcție de localizarea anatomică a cancerului din sfera ORL (împrospătarea bazei de date cu noi pacienți / date va fi un proces continuu de-a lungul proiectului).

Activitățile științifice realizate în perioada de implementare aferentă primului an de cercetare (2025) sunt în linie cu obiectivele propuse, după cum urmează:

- Efectuarea unei recenzii a literaturii de specialitate cu privire la rolul inteligenței artificiale în imagistica medicală - analiza datelor și publicarea rezultatelor sub formă de articol WoS;
- Efectuarea unei recenzii a literaturii de specialitate cu privire la rolul machine learning în radioterapia cancerelor din sfera ORL, identificarea caracteristicilor radiomice relevante cancerelor ORL și identificarea datelor clinice relevante care influențează rezultatul tratamentului - analiza datelor și publicarea rezultatelor sub formă de articol WoS;
- Efectuarea unei recenzii a literaturii de specialitate cu privire la corelațiile dintre semnătura radiomică (caracteristicile radiomice) ale cancerelor ORL și parametrii radiobiologici specifici tumorilor ORL, cu focalizare pe abilitatea de proliferare - analiza datelor și publicarea rezultatelor sub formă de articol WoS;
- Extragerea datelor pacienților din baza de date clinică (incluzând imagini CT și informații clinice specifice). Identificarea grupului de pacienți care necesită imagistica suplimentară pentru evaluarea sfârșitului de tratament;
- Anonimizarea datelor în vederea protecției datelor personale a pacienților;
- Normalizarea imaginilor pentru a asigura uniformitatea extragerii parametrilor;
- Realizarea bazei de date comasate (care se va dezvolta în continuare pe parcursul proiectului), incluzând informațiile relevante proiectului, pe subgrupe de pacienți în funcție de localizarea anatomică a cancerului din sfera ORL.

Activitățile de diseminare a rezultatelor proiectului aferente primului an de cercetare (2025) au avut loc pe mai multe căi, după cum urmează:

- Realizarea website-ului,
- Publicarea a trei articole în reviste indexate WoS cu AIS în quartile – Q1, Q2 și Q4,
- O prezentare la conferință internațională,
- O teză de doctorat în domeniul proiectului finalizată și susținută la universitatea gazdă.

De asemenea, sunt în lucru un articol științific precum și un abstract pentru o conferință internațională, care vizează fluidizarea workflow-ului radioterapiei prin utilizarea IA în segmentarea imaginilor pacienților cu cancer ORL în vederea realizării planului de tratament.

Referințe bibliografice:

1. Altinok O, et al. Phys Med. 2023;114:102671.
2. Bogowicz M, et al. Acta Oncologica. 2017a;56(11):1531-6.
3. Bogowicz M, et al. Int J Radiat Oncol Biol Phys. 2017b;99(4):921-8.
4. Bogowicz M, et al. Sci Rep 2019;9:15198.
5. Chen L, et al. Cancer Imaging. 2020;20:28.
6. Chen W, et al. BMC Cancer. 2024;24(1):418.
7. He Z, et al. Eur Radiol. 2022;32(12):8099-110.
8. Kazmierski M, et al. Cancer Res Commun. 2023;3(6):1140-51.
9. Lin N, et al. J Comput Assist Tomogr. 2022;46(5):823-9.
10. Lin M, et al. Acad Radiol. 2023;30(10):2201-2211.
11. Lin CH, et al. Radiother Oncol. 2023;189:109938.
12. Osapoetra LO, et al. Sci Rep. 2021;11(1):6117.
13. Sher DJ, et al. Clin Cancer Res. 2023;29(17):3284-91.
14. Song LL, et al. BMC Med Imaging. 2021;21(1):54.
15. Suh CH, et al. Sci Rep. 2020;10(1):17525.
16. Tomita H, et al. Eur Radiol. 2021;31(10):7440-9.
17. Wang Y, et al. Front Oncol. 2024;14:1423304.
18. Yang L, et al. Magn Reson Imaging. 2025;116:110276.
19. Zhao X, et al. Eur Radiol. 2023;33:1121-31.
20. Zheng YM, et al. Eur Radiol. 2023;33(3):2160-2170.

Director Proiect,
Marcu Loredana Gabriela

Oradea, Nov 2025