

---

REZUMATUL TEZEI DE DOCTORAT

# Profilul metabolomic al pacienților cu cancer colorectal. Implicații în stadializare, diagnostic și prognostic

---

Doctorand **Claudiu Răchieriu**

---

Conducător de doctorat Prof.dr. **Nadim Al Hajjar**

---



**UMF**  
UNIVERSITATEA DE  
MEDICINĂ ȘI FARMACIE  
IULIU HAȚIEGANU  
CLUJ-NAPOCA

# CUPRINS

<b>INTRODUCERE</b>	15
<b>STADIUL ACTUAL AL CUNOAȘTERII</b>	
<b>1. Introducere</b>	19
1.1. Epidemiologie	19
1.2. Diagnostic	20
1.3. Stadializare	21
1.4. Tratatamentul cancerului colorectal	22
1.5. Evoluție	23
1.6. Terminologie	24
1.7. Metabolomica în CCR	25
<b>2. Aplicații clinice în CCR</b>	28
2.1. Factori predispozanți	28
2.1.1. Fumatul	28
2.1.2. Alcoolul	29
2.1.3. Tulburări metabolice	29
2.2. Diagnostic precoce	30
2.3. Diferențierea între adenom și tumoră malignă	32
2.4. Stadializare	32
2.5. Evoluția CCR și relația cu chimioterapia	33
2.6. Prognostic și supraviețuire	35
2.7. Căi metabolice implicate în CCR	36
2.7.1. Metabolismul glucidic	36
2.7.2. Metabolismul lipidic (lipidomica)	37
2.7.3. Metabolismul aminoacizilor	40
2.7.4. Metabolismul nucleotidelor	42
2.7.5. Alte căi metabolice	44
2.7.5.1. Carnitina	44
2.7.5.2. Bilirubina și acizii biliari	45
2.7.5.3. Nicotinamida, NAD și derivații	46
2.7.5.4. Acidul picolinic	47
2.8. Relația dintre CCR și alte căi metabolice	48
2.8.1. Dislipidemie	48
2.8.2. Adipokinele	49
2.8.3. Semnătura metabolică a obezității	49

2.8.4. Diabetul zaharat și CCR din perspectiva metabolomică	50
<b>CONTRIBUȚIA PERSONALĂ</b>	
<b>1. Ipoteza de lucru/obiective</b>	53
<b>2. Metodologie generală</b>	55
2.1. Pacienți și conformitatea cu cerințele etice standard	55
2.2. Recoltarea și procesarea probelor de sânge	58
2.3. Analiza serului prin HPLC-ESI(+)-QTOF-MS	58
2.4. Analiza statistică	59
<b>3. Studiul 1 - Analiza metabolomică nețintită a pacienților cu CCR în vederea evidențierii de potențiali biomarkeri</b>	61
3.1. Introducere	61
3.2. Ipoteza de lucru/obiective	62
3.3. Material și metodă	63
3.3.1. Pacienți și conformitatea cu cerințele etice standard	63
3.3.2. Recoltarea și procesarea probelor de sânge	63
3.3.3. Analiza serului prin HPLC-ESI(+)-QTOF-MS	63
3.3.4. Analiza statistică	64
3.4. Rezultate	64
3.4.1. Analiza multivariată	64
3.4.1.1. Normalizarea valorilor m/z și a probelor	64
3.4.1.2. Parametrii VIP - Fold change (CRC/C) threshold și Volcano Plot	66
3.4.1.3. Testul T	68
3.4.1.4. Analiza principalelor componente (PCA) și analiza discriminării parțiale a celor mai mici pătrate - partial least squares discriminant analysis (PLSDA)	68
3.4.1.5. Dendograma euclidiană	71
3.4.1.6. Hărți termice (heatmaps) de corelație	72
3.4.1.7. Algoritmul Random Forrest (Pădure aleatorie) și valoarea sa predictivă	74
3.4.1.8. Analiza biomarkerilor	76
3.5. Discuții	81
3.5.1. Limitări legate de tehnica de identificare a metaboliților	81
3.5.2. Stabilitatea metabolomului în timp și statusul postprandial	82
3.5.3. Stabilitatea metaboliților în funcție de condițiile de păstrare ale	83

probelor și momentul recoltării	
3.5.4. Alți posibili factori de eroare	84
3.5.5. Compararea biomarkerilor obținuți cu literatura	84
3.6. Concluzii	89
<b>4. Studiul 2 - Analiza metabolomică a pacienților cu CCR în funcție de stadiul TNM</b>	91
4.1. Introducere	91
4.2. Ipoteza de lucru/obiective	91
4.3. Material și metodă	92
4.3.1. Pacienți și conformitatea cu cerințele etice standard	92
4.3.2. Recoltarea și procesarea probelor de sânge	93
4.3.3. Analiza serului prin HPLC-ESI(+)-QTOF-MS	93
4.3.4. Analiza statistică	93
4.4. Rezultate	94
4.4.1. Normalizarea datelor de variabile (valori m/z) și de probe din cele 5 grupuri	94
4.4.2. Analiza ANOVA univariată	95
4.4.3. Analiza principalelor componente (PCA)	100
4.4.4. Analiza discriminării parțiale a celor mai mici pătrate - partial least squares discriminant analysis (PLSDA)	101
4.4.5. Dendrograma (distanța euclidiană, algoritm Ward)	102
4.4.6. Hărți termice de corelație (Heatmap)	103
4.4.7. Metoda Random Forrest (pădure aleatorie)	105
4.4.8. Identificarea moleculelor implicate în CCR	107
4.4.9. Analiza statistică a subgrupurilor CCR bazată pe valorile de vârf MS	111
4.4.10. Analiza statistică clasică pentru diferențierea stadiilor	112
4.5. Discuții	114
4.6. Concluzii	119
<b>5. Studiul 3. Identificarea căilor metabolice modificate în CCR</b>	121
5.1. Introducere	121
5.2. Ipoteza de lucru/obiective	122
5.3. Material și metodă	122
5.4. Rezultate	123
5.5. Discuții	126
5.6. Concluzii	127

<b>6. Concluzii generale (sinteză)</b>	129
<b>7. Originalitatea și contribuțiile inovative ale tezei</b>	131
<b>REFERINȚE</b>	133

Cuvinte cheie: cancer colorectal, metabolomică, lipidomică, cromatografie lichidă, biomarkeri

## INTRODUCERE

Teza noastră se referă la profilul metabolomic al pacienților cu cancer colorectal. Cancerul colorectal este unul din cele mai frecvente procese neoplazice maligne (situat între primele trei ca incidență și mortalitate la ambele sexe, cu o importantă variație geografică legată mai ales de factori de mediu, în particular dieta). Stadializarea acestuia clinică și anatomopatologică (sistemul TNM) are mare importanță în stabilirea protocolului terapeutic, care este un tratament multidisciplinar. Acesta include chirurgia (aflată în poziția princeps), chimioterapia, radioterapia, precum și alte tratamente oncologice apărute mai recent - terapii biologice (anticorpi monoclonali împotriva unor ținte specifice exprimate de celulele tumorale) și imunologice. Există deja studii care fac distincție clară între amprentele metabolice ale stadiilor precoce, intermediare sau metastatice. Uneori stadializarea preoperatorie cu mijloacele existente (radiologice și endoscopice) este incorectă, în special în sensul substadializării, ceea ce poate duce la intervenții chirurgicale inutile. Acesta a fost unul din motivele principale pentru care știința încearcă să găsească metode mai precise și dacă este posibil neinvazive pentru stadializare și reprezintă și motivația cercetării acestei teze.

Ultimele două decenii au fost marcate de progresul științelor medicale prin dezvoltarea unor domenii ce fac generic parte din familia „omics” (genomică, proteomică, metabolomică etc.). Dintre acestea, *metabolomica* se ocupă de studiul moleculelor mici ce pot avea rol de substrat, produs intermediar sau final al unor reacții metabolice în organism. Acestea pot aparține diferitelor clase de molecule: glucide, aminoacizi, lipide, nucleotide, vitamine, cofactori etc. Există mai multe variante tehnice, cele mai utilizate fiind cromatografia lichidă sau gazoasă, cuplată cu spectrometria de masă. Se identifică combinații de metaboliți (amprente metabolice) prelevate din probe biologice (țesut, ser, plasmă) caracteristice unor anumite procese biologice.

În cazul medicinei cel mai important domeniu de utilitate al metabolomicii ar fi cel al oncologiei. Aplicația majoră a acestei ramuri este identificarea de biomarkeri, care să permită diagnosticul precoce (uneori pot aprecia chiar predispoziția), monitorizarea progresiei bolii sau depistarea precoce a eventualelor recidive, apariția metastazelor, predicția răspunsului la tratament oncologic. Există deja câteva studii care arată superioritatea anumitor combinații de metaboliți față

de markerii oncologici deja utilizați (pentru cancerul colorectal - antigenul carcinoembrionar ACE).

Dacă evaluarea se referă în mod special la modificările moleculelor lipidice implicate în metabolism, această metodă de screening este denumită „lipidomică”. Aceasta include două tipuri de tehnici: *lipidomica nețintită* (untargeted), o analiză cuprinzătoare a tuturor moleculelor lipidice măsurabile dintr-o probă de analizat, inclusiv substanțe chimice încă necunoscute și *lipidomica țintită*, ce măsoară grupuri definite de lipide cunoscute, ce sunt deja caracterizate și adnotate chimic. Această tehnologie folosește platforme analitice și tehnici avansate, cum ar fi cromatografia gazoasă sau cromatografia lichidă de înaltă performanță cuplată cu spectrometria de masă (GC-PS, HPLC-MS), rezonanța magnetică nucleară (NMR), electroforeza capilară (CE), cu perfecționări continue în vederea identificării de constelații, modele de metaboliți (amprente) în legătură cu căi metabolice specifice sau patologii particulare.

Partea teoretică („**Stadiul actual al cunoașterii**”) se bazează pe un minireview al echipei noastre publicat în 2020. Articolele legate de aplicațiile clinice ale metabolomicii în cancerul colorectal au fost ordonate începând cu câteva exemple legate de rolul metabolomicii în legătura cu factorii predispozanți, urmând apoi o trecere în revistă a datelor din literatură privitoare la aplicațiile clinice în legătură cu implicațiile în factorii predispozanți, diagnosticul precoce, diferențierea adenom-tumoră malignă, stadializarea cancerului colorectal, evoluția și monitorizarea progresiei bolii. Ulterior am detaliat modificările specifice ale claselor specifice de metaboliți (glucide, aminoacizi, lipide, nucleotide, alte) cu accent pe lipidomică. În final am tratat din perspectivă metabolomică relația dintre cancerul colorectal și alte boli metabolice, în speță diabetul zaharat, obezitatea și displidemiile.

## CONTRIBUȚIA PERSONALĂ

**1. Ipoteza de lucru/obiective.** Diversi autori au găsit modele metabolice particulare (amprente) caracteristice CCR, ce par specifice, cu sensibilitate mai crescută decât markerii existenți deja în urmărire postoperatorie, în diagnostic precoce, monitorizarea evoluției, stadializare, diferențierea adenom-adenocarcinom, prognostic, estimarea răspunsului la chimioterapie. Unele studii găsesc o corelație bună cu stadiul TNM, altele diferențiază trei stadii - precoce, intermediar și metastatic. Pornind de la aceste premise scopul principal al lucrării de față este identificarea de biomarkeri serici specifici din probe recoltate de la pacienți diagnosticați cu CCR (prin colonoscopie cu biopsie și examen histopatologic sau suspiciune imagistică înaltă preoperatorie, urmată de intervenție chirurgicală în condiții de urgență și confirmare histopatologică pe piesa de rezecție), aflați în diverse stadii tumorale, comparați cu voluntari fără CCR sau alte

patologii semnificative. Toată această evaluare metabolomică va fi apreciată prin tehnica de vârf a UPLC-QTOF-ESI+MS (ultra-performance liquid chromatography - quadrupole time of flight - electrospray ionization + mass spectrometry).

**2. Metodologia generală.** Lotul de pacienți cu CCR a inclus 25 de subiecți, 16 de sex masculin și 9 de sex feminin, operați în Clinica Chirurgie III din cadrul I.R.G.H. "Octavian Fodor" Cluj-Napoca, în perioada mai-decembrie 2018. Pacienții au avut toți diagnosticul de adenocarcinom colorectal confirmat prin examen anatomopatologic, fie în urma biopsiei colonoscopice efectuată în urma suspiciunii clinice, de laborator (de obicei anemie) sau imagistice, fie examinare a piesei de rezecție sau biopsiei postoperatorii (cu suspiciune radiologică preoperatorie de CCR, cu semnarea consimțământului informat preoperator). Grupul de martori a inclus 16 subiecți - 10 bărbați cu vârstă medie 53,1+/-6,7 ani și 6 femei cu vârstă medie 56,33+/-8,99, cu proporție între sexe și comorbidități asemănătoare cu grupul de pacienți.

Au fost analizate probe de ser într-un laborator specializat, prin cromatografie cu presiune ultraînaltă, cuplată cu spectrometrie de masă. Limitele de masă au fost setate între 50-1000 m/z. Au fost identificate diverse molecule prin raportul m/z, prelucrate statistic și în final cele mai semnificative molecule cu potențial de biomarkeri au fost identificate chimic cu ajutorul celor mai relevante 2 baze de date: LIPID MAPS Lipidomics Gateway și Human Metabolomics Database.

### **3. Studiu 1. Analiza metabolomică neștințită a pacienților cu CCR în vederea evidențierii de potențiali biomarkeri**

**Introducere și ipoteza de lucru.** Metabolomica neștințită este o analiză comprehensivă a tuturor moleculelor mici măsurabile dintr-o probă de lichid biologic. Se pornește de la ideea că analiza bazată pe tehnici metabolomice neștințite prin spectrometrie de masă (MS) va da naștere unui număr mare de molecule mici "identificate". Datele pot fi utilizate pentru o relativă cuantificare în rândul grupurilor de probe analizate și pentru a genera ipoteze care pot fi ulterior studiate prin abordări țintite. Obiectivul acestui studiu este menționat în titlu și constă în identificarea de posibili biomarkeri prin utilizarea tehnicii de metabolomică neștințită a unui lot de pacienți cu CCR aflați în diverse stadii evolutive, comparat cu lotul martor. Potrivit literaturii ne așteptăm să obținem modificări la nivelul căilor metabolice ale glicolizei (preferată de celula tumorală în dauna fosforilării oxidative - „efect Warburg”) și fosforilării mitocondriale (efect Warburg inversat - descris mai recent), metabolizării glutaminei exogene și mai ales ale căilor metabolice lipidice. Dintre acestea din urmă menționăm sinteza, desaturarea, elongarea de acizi grași și oxidarea mitocondrială. Modificările lipidelor în CCR pot fi explicate prin rata de proliferare accelerată a celulelor CCR, cu nevoi energetice crescute, modificări ale nivelurilor serice ale componențelor fosfolipidice rezultate din degradarea membranelor celulare (în special metaboliții colinei), precum și

modificări inflamatorii, cu modificări ale metaboliților acidului arahidonic la nivel seric sau tisular.

**Material și metodă.** Lotul de pacienți și cel al martorilor descrise mai sus au furnizat probe de ser ce au fost analizate prin cromatografie lichidă cu spectrometrie de masă (vezi mai sus). După alinieri succesive și normalizarea datelor din matrice s-a efectuat analiza multivariată. Aceasta a constatat în reprezentări de tip fold change (raportul dintre valorile medii ale unui parametru obținut în lotul de control comparativ cu lotul de examinat), grafice scatter de tip volcano plot, analiza componentelor principale - principal component analysis (PCA), analiza discriminării parțiale a celor mai mici pătrate - partial least squares discriminant analysis (PLSDA) și metode de regresie de tip random forest, descoperirea corelațiilor între probe și între variabile (valorile m/z), precum și construirea de hărți termice („heatmap”) care să reprezinte corelația dintre variabile și probe. La final, s-au generat curbele de funcționare ale receptorului (the receiver operating curves - ROCs) și valorile ariilor de sub curbele ROC (AUCs) iar apoi moleculele identificate au fost stratificate în funcție de sensibilitatea/specificitatea lor. Identificarea moleculelor care pot fi luate în considerare ca și potențiali biomarkeri a fost făcută cu ajutorul celor mai relevante 2 baze de date: LIPID MAPS Lipidomics Gateway și Human Metabolomics Database.

**Rezultate.** Prin metodele statistice menționate mai sus au fost identificate primele 25 de molecule după semnificația statistică a curbelor ROC. Acestea fac parte din familiile ceramidelor, sfingomielinelor, fosfatidil- și lizofosfatidilolinolinelor, mono- și diacilglicerolilor, acidul taurocolic, 1-OH-vitamina D, all-trans -retiniloleat, 2 esteri grași (linoleil stearat și stearil palmitat )și fosfatidilglicerolul. Ultimele 3 molecule nu au mai fost descrise anterior în literatură în această patologie.

**Concluzii.** În mare parte datele noastre sunt concordante cu cele regăsite și în alte studii, unele sunt inconstante în literatură iar altele nu se regăsesc în cercetări anterioare. Ultima afirmație se referă la esteri grași și la fosfatidilglicerol, a căror creștere are însă explicații plauzibile.

Majoritatea claselor menționate la rezultate au legătură cu degradarea membranelor celulare și cu accentuarea metabolismului celulei tumorale, altele pot fi explicate prin factori alimentari, legați de dietă, cunoscuți ca predispozanți pentru patologia tumorală în discuție. Nu am regăsit între cei mai semnificativi 25 biomarkeri metaboliți ai acidului arahidonic, așa cum anticipasem în ipoteza de lucru.

#### **4. Studiu 2. Analiza metabolomică a pacienților cu CCR în funcție de stadiul TNM**

**Introducere și ipoteza de lucru.** Bazându-ne pe datele din literatură și pe studiul nostru anterior vom compara din punct de vedere metabolomic probele serice ale pacienților cu CCR grupați în stadii TNM atât cu grupul de control cât și



între stadii. Separarea pe stadii TNM s-ar putea să fie certă sau să respecte unele modele obținute în stadii anterioare în care stadiile precoce (I și II) au valori apropiate, net diferite de stadiile avansate (III și mai ales IV-cu metastaze hepatice). Nu vom putea evalua ipoteza potrivit căreia profilurile metabolomice ale pacienților cu metastaze hepatice sunt diferite de cele ale pacienților cu metastaze extrahepatice, ultimul grup nefiind reprezentat în lotul nostru de studiu. O altă posibilitate este ca unele variabile să aibă valori care să aibă tendințe modificate continuu (fie în sens crescător fie descrescător) între cele patru stadii.

**Material și metodă.** Împărțirea pe stadii TNM a generat 3 pacienți în stadiul I, 8 pacienți în stadiul II, 6 pacienți în stadiul III și 8 pacienți în stadiul IV. Pacienții în stadiul IV au avut metastaze exclusiv hepatice. Aceste loturi au fost comparate între ele și cu lotul martor. Analiza statistică a inclus analiza ANOVA univariată, cu calculul coeficienților de corelație între stadii și în raport cu grupul de control. După alinieri succesive și normalizarea datelor din matrice s-a efectuat analiza multivariată. Aceasta a constat în reprezentări de tip fold change (raportul dintre valorile medii ale unui parametru obținut în lotul de control comparativ cu lotul de examinat), grafice scatter de tip volcano plot, analiza componentelor principale - principal component analysis (PCA), analiza discriminării parțiale a celor mai mici pătrate - partial least squares discriminant analysis (PLSDA) și metode de regresie de tip random forest, descoperirea corelațiilor între probe și între variabile (valorile m/z), precum și construirea de hărți termice („heatmap”). Identificarea moleculelor s-a făcut similar cu studiul 1.

**Rezultate.** Ca și în cazul comparării globale a pacienților cu CCR cu martorii, în diferențierea pe stadii, se identifică biomarkeri predominant lipidici, ce pot fi considerați markeri ai progresiei bolii, care fac parte în special din familia fosfatidil- și lizofosfatidil colinelor, ceramidelor, sfingomielinelor, mono- și diacilglicerolilor, esterilor de colesterol, metaboliților vitaminelor liposolubile A și D. Remarcăm însă proporția diferită a acestor componente, cu mai puține ceramide și mult mai multe fosfatidil coline, care reprezintă circa un sfert din cele mai semnificative 25 de molecule.

În plus față de analiza din studiul I se remarcă și prezența de carnitine acilate, o mare varietate de acizi grași liberi și esterificați, metaboliți ai acidului arahidonic, estrone, acizi biliari, precum și metaboliți ai unor aminoacizi și amine (ex: acilserotonina, N-docosahexaenoyl glutamic acid, N-Stearoil fenilalanina), vitamina B2 (riboflavina), coenzima Q10 (ubiquinona).

**Concluzii.** Datele din analiza multivariată dar și cele obținute prin compararea mediilor valorilor de vârf obținute prin MS între diverse stadii arată per ansamblu diversele molecule menționate mai sus care pot fi considerați biomarkeri pentru diagnostic precoce dar și markeri de progresie ai CCR. Ca fenomen caracteristic se remarcă activarea biosintezei de acizi grași, mai ales al celor polinesaturați, și legarea lor la diferiți metaboliți (aminoacizi, amine și derivați ai fosfatidilcolinei). Per ansamblu metabolismul lipidic este activat

(upregulated) în CCR. Din punct de vedere metabolomic, în compararea dintre stadii a intensităților medii cea mai mare diferență se remarcă între stadiile III și IV, lucru care se regăsește și în unele date din literatură, unii autori afirmând că tumorile ce au capacitatea de metastazare sunt biologic diferite de cele care au evoluție locoregională. Această modificare în biologia tumorală referitoare la metastazarea hepatică par să fie dependente și de sediul metastazării.

## 5. Studiu 3. Identificarea căilor metabolice modificate în CCR

**Ipoteza de lucru.** Utilizând probele serice recoltate ale pacienților cu CCR și ale martorilor, vom încerca identificarea căilor metabolice alterate în CCR, prin analiza UPLC-QTOF-ESI+MS și ulterior utilizarea softului specializat. Ulterior vom studia în aceeași manieră ce căi metabolice pot fi identificate în progresia CCR, datele furnizate în literatură punând accent pe trecerea în stadiul IV, metastatic, unde par a fi implicate diverse kinaze, factori de semnalizare ai transcripției proteice nucleare, TNF, interleukine, dar și marea familie a ceramidelor, cu rol central în calea metabolismului sfingolipidelor.

**Material și metodă.** Am utilizat algoritmul dedicat GSEA din cadrul programului Metaboanalyst 5.0. Software-ul dedicat a generat căile metabolice alterate în CCR din bazele de date obținute de pe web și din bibliotecile de metaboliți. Într-o manieră similară am generat ulterior căile metabolice identificate în analiza diferențierii între stadii CCR și respectiv lotul C (control/martori).

**Rezultate.** Căile metabolice alterate cu rol în diagnosticul CCR au fost: naveta carnitinei, metabolismul pirimidinelor, biosinteza de novo a acizilor grași, activarea acizilor grași, metabolismul vitaminei D3, metabolismul glicerofosfolipidelor. În progresia CCR (analiza pe stadii) apar în plus metabolismul acidului arahidonic și al leucotrienelor.

**Concluzii.** Căile metabolice identificate au rol atât în constituirea membranelor celulare precum și în semnalizarea intercelulară. Căile ce apar în plus alterate în compararea pe stadii TNM (rol în progresia bolii), respectiv metabolismul acidului arahidonic și al leucotrienelor, sunt legate probabil în special de micromediul tumoral, caracterizat de un nivel de inflamație și dereglarea răspunsului imun. Acești mediatorii ai inflamației par să fie clar corelați cu progresia bolii.

## 7. Originalitatea și contribuțiile inovative ale tezei

Este primul studiu în arealul nostru geografic cu rezultate publicate în acest domeniu (România sau țările vecine) referitoare la metabolomica în cancerul colorectal. În studiul de metabolomică neîntită am identificat ca potențiali biomarkeri esteri grași (linoleil stearat și stearil palmitat) și fosfatidilglicerolul, ce nu au mai fost anterior menționați în literatură ca fiind implicați în CCR și pot constitui un punct de pornire pentru o cercetare ulterioară de metabolomică țintită.

---

PHD THESIS ABSTRACT

# Metabolomic profile of colorectal cancer patients. Implications in staging, diagnosis and prognosis

---

PhD Student **Claudiu Răchieriu**

---

PhD Scientific Coordinator Prof.dr. **Nadim Al Hajjar**

---



**UMF**  
UNIVERSITATEA DE  
MEDICINĂ ȘI FARMACIE  
IULIU HAȚIEGANU  
CLUJ-NAPOCA

# TABLE OF CONTENTS

<b>INTRODUCTION</b>	15
<b>CURRENT STATE OF KNOWLEDGE</b>	
<b>1. Introduction</b>	19
1.1. Epidemiology	19
1.2. Diagnosis	20
1.3. Staging	21
1.4. Colorectal cancer treatment	22
1.5. Evolution	23
1.6. Terminology	24
1.7. Metabolomics in CRC	25
<b>2. Clinical applications in CRC</b>	28
2.1. Predisposing factors	28
2.1.1. Smoking	28
2.1.2. Alcohol	29
2.1.3. Metabolic disorders	29
2.2. Early diagnosis	30
2.3. Differentiation between adenoma and malignant tumor	32
2.4. Staging	32
2.5. CRC evolution and relationship with chemotherapy	33
2.6. Prognosis and survival	35
2.7. Metabolic pathways altered in CRC	36
2.7.1. Carbohydrate metabolism	36
2.7.2. Lipid metabolism (lipidomics)	37
2.7.3. Aminoacid metabolism	40
2.7.4. Nucleotide metabolism	42
2.7.5. Other metabolic pathways	44
2.7.5.1. Carnitine	44
2.7.5.2. Bilirubin and bile acids	45
2.7.5.3. Nicotinamide, NAD and derivatives	46
2.7.5.4. Picolinic acid	47
2.8. The link between CRC and other metabolic disorders	48
2.8.1. Dyslipidemia	48
2.8.2. Adipokines	49
2.8.3. The metabolic signature of obesity	49

2.8.4. Diabetes and CRC from a metabolomics perspective	50
<b>PERSONAL CONTRIBUTION</b>	
<b>1. Work hypothesis/objectives</b>	53
<b>2. General method</b>	55
2.1. Patients and compliance with standard ethical requirements	55
2.2. Collection and processing of blood samples	58
2.3. Serum analysis through HPLC-ESI(+)-QTOF-MS	58
2.4. Statistical analysis	59
<b>3. Study 1 - Untargeted metabolomics analysis of patients with CRC in order to find potential biomarkers</b>	61
3.1. Introduction	61
3.2. Work hypothesis/objectives	62
3.3. Material and method	63
3.3.1. Patients and compliance with standard ethical requirements	63
3.3.2. Collection and processing of blood samples	63
3.3.3. Serum analysis through HPLC-ESI(+)-QTOF-MS	63
3.3.4. Statistical analysis	64
3.4. Results	64
3.4.1. Multivariate analysis	64
3.4.1.1. Normalization of m / z values and samples	64
3.4.1.2. VIP parameters - Fold change (CRC / C) threshold and Volcano Plot	66
3.4.1.3. T-test	68
3.4.1.4. Principal components analysis (PCA) and partial least squares discriminant analysis (PLSDA)	68
3.4.1.5. Euclidean dendrogram	71
3.4.1.6. Correlation heatmaps	72
3.4.1.7. Random Forrest algorithm and its predictive value	74
3.4.1.8. Biomarker analysis	76
3.5. Discussions	81
3.5.1. Limitations related to the technique of identification of metabolites	81
3.5.2. Metabolic stability over time and postprandial status	82
3.5.3. Stability of metabolites according to sample storage conditions and timing of collection	83
3.5.4. Other possible error factors	84

3.5.5. Comparison of obtained biomarkers with literature data	84
3.6. Conclusions	89
<b>4. Study 2 - Metabolomic analysis of patients with CRC according to TNM stage</b>	91
4.1. Introduction	91
4.2. Work hypothesis/objectives	91
4.3. Material and method	92
4.3.1. Patients and compliance with standard ethical requirements	92
4.3.2. Collection and processing of blood samples	93
4.3.3. Serum analysis through HPLC-ESI(+)-QTOF-MS	93
4.3.4. Statistical analysis	93
4.4. Results	94
4.4.1. Normalization of variable data (m / z values) and samples from the 5 groups	94
4.4.2. Univariate ANOVA analysis	95
4.4.3. Principal Component Analysis (PCA)	100
4.4.4. Partial least squares discriminant analysis (PLSDA)	101
4.4.5. Dendrogram (Euclidean distance, Ward algorithm)	102
4.4.6. Correlation heatmaps	103
4.4.7. Random Forrest method	105
4.4.8. Identification of molecules involved in RCC	107
4.4.9. Statistical analysis of CRC subgroups based on peak MS values	111
4.4.10. Classical statistical analysis for stage differentiation	112
4.5. Discussions	114
4.6. Conclusions	119
<b>5. Study 3. Identification of altered metabolic pathways in CRC</b>	121
5.1. Introduction	121
5.2. Work hypothesis/objectives	122
5.3. Material and method	122
5.4. Results	123
5.5. Discussions	126
5.6. Conclusions	127
<b>6. General conclusions (summary)</b>	129
<b>7. Originality and innovative contributions of the thesis</b>	131
<b>REFERENCES</b>	133

Keywords: colorectal cancer, metabolomics, lipidomics, liquid chromatography, biomarkers

## INTRODUCTION

Our thesis refers to the metabolomic profile of patients with colorectal cancer. Colorectal cancer is one of the most common malignant neoplastic processes (ranked among the top three in incidence and mortality in both sexes, with significant geographical variation related mainly to environmental factors, mainly diet). Its clinical and anatomopathological staging (TNM system) is of great importance in establishing the therapeutic protocol, which is a multidisciplinary treatment. This includes surgery (in the princeps position), chemotherapy, radiation therapy, as well as other more recent oncological treatments - biological therapies (monoclonal antibodies against specific targets expressed by tumor cells) and immunological treatment. There are already studies that make a clear distinction between the metabolic fingerprints of the early, intermediate or metastatic stages. Sometimes preoperative staging with existing means (radiological and endoscopic) is incorrect, especially in the sense of substadiation, which can lead to useless surgery. This was one of the main reasons why science tries to find more precise and if possible non-invasive methods for staging and is also the motivation for the research of this thesis.

The last two decades have been marked by the progress of medical sciences by developing fields that are generically part of the "omics" family (genomics, proteomics, metabolomics, etc.). Of these, metabolomics deals with the study of small molecules that can act as a substrate, intermediate or final product of metabolic processes in the organism. They can belong to different classes of molecules: carbohydrates, amino acids, lipids, nucleotides, vitamins, cofactors, etc. There are several technical variants, the most used being liquid or gas chromatography, coupled with mass spectrometry. Clusters of metabolites (metabolic fingerprints) taken from biological samples (tissue, serum, plasma) characteristic of certain biological processes are identified.

In the field of medicine, the most important use of metabolomics would be in oncology. The major application of this branch is the identification of biomarkers, which allow early diagnosis (sometimes they can even appreciate the predisposition), monitoring the progression of the disease or early detection of possible recurrences, metastases, prediction of response to cancer treatment. There are already several studies showing the superiority of certain combinations of metabolites over oncological markers already used (for colorectal cancer - CEA carcinoembryonic antigen).

If the evaluation refers specifically to changes in lipid molecules involved in metabolism, this screening method is called "lipidomics". This includes two types of

techniques: untargeted lipidomics, a comprehensive analysis of all measurable lipid molecules in a test sample, including as yet unknown chemicals, and targeted lipidomics, which measure defined groups of known lipids that are already characterized and chemically annotated. This technology uses advanced analytical platforms and techniques, such as gas chromatography or high performance liquid chromatography coupled with mass spectrometry (GC-PS, HPLC-MS), nuclear magnetic resonance (NMR), capillary electrophoresis (CE), with continuous improvements in order to identify constellations, patterns of metabolites (fingerprints) linked to specific metabolic pathways or particular pathologies.

The theoretical part ("Current state of knowledge") is based on a mini-review of our team published in 2020. Articles related to the clinical applications of metabolomics in colorectal cancer were ordered starting with some examples related to the role of metabolomics in relation to predisposing factors, followed by a review of data from the literature on clinical applications related to implications in predisposing factors, early diagnosis, adenoma-malignant tumor differentiation, colorectal cancer staging, evolution and monitoring of disease progression. Later we detailed the specific changes of specific classes of metabolites (carbohydrates, amino acids, lipids, nucleotides, others) with emphasis on lipidomics. Finally, we approached from a metabolomic perspective the relationship between colorectal cancer and other metabolic diseases, in this case diabetes, obesity and dyslipidemia.

## **PERSONAL CONTRIBUTION**

**1. Work hypothesis/objectives.** Several authors have found particular metabolic patterns (fingerprints) characteristic of RCC, which seem specific, with higher sensitivity than existing markers in postoperative follow-up, early diagnosis, monitoring progression, staging, adenoma-adenocarcinoma differentiation, prognosis, estimation of response to chemotherapy. Some studies find a good correlation with the TNM stage, others differentiate three stages - early, intermediate and metastatic. Based on these premises, the main purpose of this paper is to identify specific serum biomarkers from samples collected from patients diagnosed with CRC (confirmed by colonoscopy with biopsy and histopathological examination or high preoperative imaging suspicion, followed by emergency surgery and confirmation histopathological on the resection specimen), in various tumor stages, compared with volunteers without RCC or other significant pathologies. All this metabolomic evaluation will be assessed by the state-of-the-art UPLC-QTOF-ESI + MS (ultra-performance liquid chromatography - quadrupole time of flight - electrospray ionization + mass spectrometry).

**2. General method.** The group of patients with RCC included 25 subjects, 16 males and 9 females, operated in the Surgery Clinic III within the I.R.G.H. "Octavian Fodor" Cluj-Napoca, between May and December 2018. All patients had a diagnosis



of colorectal adenocarcinoma confirmed by anatomopathological examination, either following a colonoscopic biopsy performed following clinical, laboratory (usually anemia) or imaging suspicion, or examination of the resection specimen or postoperative biopsy (with preoperative radiological suspicion of RCC, with the signing of the preoperative informed consent). The control group included 16 subjects - 10 men with a mean age of 53.1 +/- 6.7 years and 6 women with a mean age of 56.33 +/- 8.99, with a similar proportion of sexes and comorbidities to the patient group.

Serum samples were analyzed in a specialized laboratory by ultra-high pressure chromatography, coupled with mass spectrometry. The mass limits were set between 50-1000 m/z. Various molecules were identified by the m / z ratio, statistically processed and finally the most significant molecules with biomarker potential were chemically identified using the 2 most relevant databases: LIPID MAPS Lipidomics Gateway and Human Metabolomics Database.

### **3. Study 1. Untargeted metabolomics analysis of patients with CRC in order to find potential biomarkers**

**Introduction and working hypothesis.** Untargeted metabolomics is a comprehensive analysis of all measurable small molecules in a biological fluid sample. It starts from the idea that analysis based on untargeted metabolomics techniques by mass spectrometry (MS) will result in a large number of small "identified" molecules. The data can be used for a relative quantification among the groups of samples analyzed and to generate hypotheses that can be subsequently studied by targeted approaches. The objective of this study is mentioned in the title and consists in identifying possible biomarkers by using the technique of untargeted metabolomics on a group of samples collected from patients with CRC in various TNM stages, compared to the control group. According to the literature, we expect to obtain changes in the metabolic pathways of glycolysis (preferred by the tumor cell to the detriment of oxidative phosphorylation - "Warburg effect") and mitochondrial phosphorylation (inverted Warburg effect - described more recently), exogenous glutamine metabolism and especially lipid metabolic pathways. Among the latter we mention the synthesis, desaturation, elongation of fatty acids and mitochondrial oxidation. Changes in lipids in CCR can be explained by the rate of accelerated proliferation of CCR cells, with increased energy needs, changes in serum levels of phospholipid components resulting from the degradation of cell membranes (especially choline metabolites), and inflammatory changes, with changes in arachidonic acid metabolites in serum or tissue.

**Material and method.** The patient group and the control group described above provided serum samples which were analyzed by liquid chromatography with mass spectrometry (see above). After successive alignments and normalization of data from the matrix, multivariate analysis was performed. This

consisted of fold change representations (the ratio between the average values of a parameter obtained in the control group compared to the group to be examined), volcano plot scatter graphs, principal component analysis (PCA), partial least squares discriminant analysis (PLSDA) and random forest regression methods, the discovery of correlations between samples and between variables (m / z values), as well as the construction of heatmaps to represent the correlation between variables and samples. Finally, the receiver operating curves (ROCs) and the values of the areas under the ROC curves (AUCs) were generated and then the identified molecules were stratified according to their sensitivity / specificity. The identification of molecules that can be considered as potential biomarkers was made using the 2 most relevant databases: LIPID MAPS Lipidomics Gateway and Human Metabolomics Database.

**Results.** Using the statistical methods mentioned above, the first 25 molecules were identified according to the statistical significance of the ROC curves. These belong to the families of ceramides, sphingomyelins, phosphatidyl- and lysophosphatidylcholines, mono- and diacylglycerols, taurocholic acid, 1-OH-vitamin D, all-trans-retinyl-oleate, 2 fatty esters (linoleyl stearate and stearyl palmitate) and phosphatidylglycerol. The last 3 molecules have not been previously described in the literature as related to this pathology.

**Conclusions.** Most of our data are consistent with those found in other studies, some are inconsistent in the literature and others are not found in previous research. The last statement refers to fatty esters and phosphatidylglycerol, the increase of which, however, has plausible explanations.

Most of the classes mentioned in the results are related to the degradation of cell membranes and the increase in tumor cell metabolism, others can be explained by dietary factors, known as predisposing to the tumor pathology in question. We did not find among the most significant 25 metabolites biomarkers of arachidonic acid, as I had anticipated in the working hypothesis.

#### **4. Study 2. Metabolomic analysis of patients with CRC according to TNM stage**

**Introduction and working hypothesis.** Based on data from the literature and our previous study we will metabolically compare the serum samples of patients with CRC grouped in TNM stages both with the control group and between stages. The separation on TNM stages may be certain or follow some models obtained in previous stages in which the early stages (I and II) have close values, clearly different from the advanced stages (III<sup>rd</sup> stage and especially liver metastases IV<sup>th</sup> stage). We will not be able to evaluate the hypothesis according to which the metabolomic profiles of patients with liver metastases are different from those of patients with extrahepatic metastases, the last group not being represented in our study group. Another possibility is that some variables have

values that have continuously changing trends (either up or down) between the four stages.

**Material and method.** The division into TNM stages generated 3 patients in stage I, 8 patients in stage II, 6 patients in stage III and 8 patients in stage IV. Stage IV patients had exclusively liver metastases. These groups were compared with each other and with the control group. The statistical analysis included the univariate ANOVA analysis, with the calculation of the correlation coefficients between stages and in relation to the control group. After successive alignments and normalization of data from the matrix, multivariate analysis was performed. This consisted of fold change representations (the ratio between the average values of a parameter obtained in the control group compared to the group to be examined), volcano plot scatter graphs, principal component analysis (PCA), partial least squares discriminant analysis (PLSDA) and random forest regression methods, the discovery of correlations between samples and between variables ( $m/z$  values), as well as the construction of heatmaps. The identification of the molecules was done similarly to study 1.

**Results.** As in the case of the global comparison of CRC patients with controls, in the differentiation by stages, lipid biomarkers are predominantly identified and can be considered markers of disease progression, which belong especially to the phosphatidyl- and lysophosphatidyl, choline, ceramide, sphingomyelin families, mono- and diacylglycerols, cholesterol esters, metabolites of fat-soluble vitamins A and D. However, we note the different proportion of these components, with fewer ceramides and many more phosphatidyl cholines, which represent about a quarter of the most significant 25 molecules.

In addition to the analysis in study I, we found the presence of acylated carnitines, a wide variety of free and esterified fatty acids, metabolites of arachidonic acid, estrones, bile acids, as well as metabolites of some amino acids and amines (eg acylserotonin, N -docosahexaenoyl glutamic acid, N-Stearoyl phenylalalanine), vitamin B2 (riboflavin), coenzyme Q10 (ubiquinone).

**Conclusions.** The data from the multivariate analysis but also those obtained by comparing the averages of the MS peak values between various stages show as a whole the various molecules mentioned above that can be considered biomarkers for early diagnosis but also progression markers of CRC. A characteristic phenomenon is the activation of fatty acid biosynthesis, especially polyunsaturated ones, and their binding to various metabolites (amino acids, amines and phosphatidylcholine derivatives). Overall, lipid metabolism is upregulated in CCR. From a metabolomic point of view, in the comparison of average intensities between stages, the biggest difference is between stages III and IV, which is also found in some data in the literature, some authors stating that tumors that have the ability to metastasize are biologically different from those with locoregional evolution. This change in tumor biology related to liver metastasis also appears to be dependent on the site of metastasis.

## 5. Study 3. Identification of altered metabolic pathways in CRC

**Work hypothesis.** Using serum samples collected from patients with CRC and controls, we try to identify altered metabolic pathways in CRC, by UPLC-QTOF-ESI + MS analysis and then using specialized software. Later we study in the same manner which metabolic pathways can be identified in the progression of CRC, the data provided in the literature emphasising the transition to stage IV, metastatic, where various kinases, signaling factors of nuclear protein transcription, TNF, interleukins, but also the large family of ceramides, with a central role in the metabolism of sphingolipids, seem to be involved.

**Material and method.** We used the dedicated GSEA algorithm within the Metaboanalyst 5.0 program. Dedicated software generated altered metabolic pathways in CRC from web-based databases and metabolite libraries. In a similar manner, we subsequently generated the metabolic pathways identified in the analysis of the differentiation between CCR stages and group C (control / controls), respectively.

**Results.** Altered metabolic pathways with a role in the diagnosis of RCC were: carnitine shuttle, pyrimidine metabolism, de novo fatty acid biosynthesis, fatty acid activation, vitamin D3 metabolism, glycerophospholipid metabolism. In the progression of CRC (staged analysis) the metabolism of arachidonic acid and leukotrienes appear in addition.

**Conclusions.** The identified metabolic pathways have a role both in the formation of cell membranes and in intercellular signaling. The pathways that appear to be altered when comparing TNM stages (role in disease progression), respectively the metabolism of arachidonic acid and leukotrienes, are probably linked especially to the tumor microenvironment, characterized by a level of inflammation and immune response disorder. These mediators of inflammation appear to be clearly correlated with disease progression.

## 7. Originality and innovative contributions of the thesis

It is the first study in our geographical area (Romania or neighboring countries) with results published in this field regarding metabolomics in colorectal cancer. In the untargeted metabolomics study we identified as potential biomarkers fatty esters (linoleyl stearate and stearyl palmitate) and phosphatidylglycerol, which have not been previously mentioned in the literature as being involved in CRC and may be a starting point for further targeted metabolomics research.