

UNIVERSITATEA BABEȘ-BOLYAI CLUJ-NAPOCA
FACULTATEA DE FIZICĂ
SPECIALIZAREA BIOFIZICĂ ȘI FIZICĂ MEDICALĂ



LUCRARE DE DIZERTAȚIE

Îndrumător

Prof. dr. Nicolae LEOPOLD

Absolvent

Ramona-Gabriela COZAN

2022

UNIVERSITATEA BABEȘ-BOLYAI CLUJ-NAPOCA
FACULTATEA DE FIZICĂ
SPECIALIZAREA BIOFIZICĂ ȘI FIZICĂ MEDICALĂ



Biopsia lichidă SERS a biofluidelor în diagnosticarea cancerului de sân

Lucrare de dizertație

Îndrumător

Prof. dr. Nicolae LEOPOLD

Absolvent

Ramona-Gabriela COZAN

2022

ABSTRACT

There is a significant need for new techniques in medical screening that do not require specialist operation and have a high level of accessibility and accuracy. Surface-enhanced Raman scattering (SERS) is the spectroscopic analysis that fits the criteria previously mentioned and provides the metabolic composition of biofluids in order to characterize complex biological systems for diagnostic and therapeutic purposes.

This work aims to analyse the molecular and metabolic information provided by label-free SERS liquid biopsy on serum and urine samples from breast cancer patients and controls and to compare the performance of five different machine learning (ML) algorithms applied to the SERS data.

The study shall present the applicability of SERS spectroscopy to breast cancer diagnosis. The diagnostic technique on which it is based is a simple, minimally invasive biopsy, analysis being performed on patient filtered blood serum and urine. This research involves 60 serum and urine samples obtained from the same patients, of which 39 samples were patients with breast cancer (BC) and 21 samples control participants (CTRL). All the samples were analyzed by SERS spectroscopy. Afterwards, Decision Tree (DT), Random Forest (RF), Gaussian Naive Bayes (GNB), Support Vector Machine (SVM) and Linear Discriminant Analysis (LDA) were applied to the SERS spectra to classify the BC and CTRL patients based on serum and urine samples.

The SERS spectra of serum and urine samples show high similarities, both being dominated by purine metabolite SERS bands (uric acid, hypoxanthine and xanthine). Besides purine metabolite SERS bands, urine samples contain SERS bands of creatinine and urea as well. Further, to classify the BC and CTRL patients we implement the ML algorithms and found similar classification accuracies in the range of 61-89%, regardless the ML model implied. LDA performed the best classification accuracy for serum spectra (83%), while DT show the best performance for urine spectra (89%). By concatenating the urine and serum SERS data DT and RF assessed the best classification accuracy (83%).

This research serves as a benchmark in order to improve the correlation of SERS bands with metabolic and biochemical changes in future SERS-based studies and supports the possibilities and advantages of the applicability of the SERS spectroscopy in this field.

CUPRINS

INTRODUCERE	4
CAPITOLUL 1.	
SPECTROSCOPIA RAMAN AMPLIFICATĂ DE SUPRAFAȚĂ (SERS)	6
1.1. Abordare teoretică	6
1.2. SERS în medicină	12
CAPITOLUL 2.	
CANCERUL DE SÂN (MAMAR)	16
CAPITOLUL 3.	20
MATERIALE ȘI METODE	
3.1. Colectarea, pregătirea și depozitarea probelor de ser și urină	20
3.2. Substratul SERS	20
3.3. Metodologia SERS utilizată	22
3.4. Pre-procesarea și analizarea spectrelor SERS	22
3.5. Algoritmii de machine learning (ML)	24
CAPITOLUL 4.	
REZULTATE ȘI DISCUȚII	27
4.1. Analiza comparativă și complementară a informațiilor metabolice furnizate de spectrelor SERS ale fracțiunii metabolice din ser filtrat și urinei	27
4.2. Clasificarea spectrelor SERS în canceros vs necanceros prin algoritmi de machine learning (ML)	33
CONCLUZII	37
BIBLIOGRAFIE	38

INTRODUCERE

În ultimii ani s-a remarcat o nevoie din ce în ce mai mare de noi abordări în screening-ul sau diagnosticul medical, tehnici cu un grad ridicat de accesibilitate, sensibilitate și specificitate. În prezent, studiile se concentrează pe tehnici rapide, minim invazive și ușor de utilizat, capabile de a identifica timpuriu boli, iar spectroscopia Raman amplificată de suprafață (*surface-enhanced Raman scattering* – SERS) a câștigat cea mai mare atenție. SERS este un tip de spectroscopie vibrațională, care oferă informații metabolice și moleculare ale probei, bazate pe excitarea nivelurilor de energie vibrațională de către fotonii împrăștiați inelastic. Pe lângă specificitatea moleculară ridicată, SERS oferă, de asemenea, o sensibilitate ridicată, atingând chiar detecția unei singure molecule [1, 2].

Biopsia lichidă SERS este o abordare în care algoritmi de machine learning (ML) permit clasificarea probelor în pacienți și control pe baza semnalului SERS al biofluidelor și a demonstrat progrese semnificative în ultimul deceniu [3-5]. Au fost raportate atât studii SERS pe celule, pentru determinarea distribuției celulare a unor compuși chimici sau pentru urmărirea metabolismului celular [6], cât și studii pe fluide biologice precum ser sagvin, plasmă, urină, lichid amniotic sau salivă pentru diagnosticarea rapid și minim invazivă într-o multitudine de tipuri de cancer [4, 7-10], boli autoimune [10, 11] și degenerative [12], noduli tiroidieni benigni [13], sau chiar boli coronariene [14]. Dintre diferitele abordări SERS pentru diagnosticul medical (de exemplu, analiza SERS a țesuturilor), biopsia lichidă SERS a fost până acum cea mai populară alegere, deoarece este minim invazivă, necesită volume mici de probă și este o tehnică rapidă de achiziție.

Precizia generală a diagnosticului medical bazat pe SERS atinge la momentul actual o valoare de ~80–90% [7, 15-18]. Deși această valoare rămâne sub unele tehnici standard de laborator clinic, ea evidențiază potențialul SERS ca instrument versatil în cadrul point-of-care, fiind o metodă de screening rapidă și ușor de utilizat.

În ciuda numărului tot mai mare de articole de cercetare în domeniul biopsiei lichide bazate pe SERS pentru diagnosticul medical, există încă multe îmbunătățiri ce pot fi aduse. Unul dintre deficiențele biopsiei lichide SERS este lipsa de atribuire precisă a benzilor SERS, analiză ce prezintă potențialul de a identifica variații în compoziția metabolică a biofluidelor pacienților.

Prin urmare, lucrarea de față urmărește să analizeze sistematic informațiile metabolice extrase din biopsia lichidă SERS din probe de ser filtrat și deproteinizat și urină. Spectrele SERS ale serului filtrat și urinei au fost comparate cu spectrele SERS ale metaboliților purinici, pentru a evidenția contribuția majoră a acestora la spectrele SERS ale celor două biofluide. Pe lângă atribuirea benzilor SERS ale probelor de ser filtrat și deproteinizat și urină, un al doilea scop la fel de important a fost evaluarea performanței relative de clasificare a cinci modele diferite de machine learning pentru clasificarea cancerului de sân și a probelor de control, pe baza spectrelor SERS de fracțiune metabolică a serului și urină.

Rezultatele studiului au fost diseminate în articolul științific **SERS liquid biopsy in breast cancer. What can we learn from SERS on serum and urine?** din revista **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, Volumul 273, 15 Mai, 2022 (doi.org/10.1016/j.saa.2022.120992) și în cadrul conferinței internaționale **Metabolomics 2022**.

Nu în ultimul rând, doresc să-i mulțumesc doctorandei Ștefania Iancu pentru ajutorul acordat în realizarea acestui studiu.

CAPITOLUL 1.

SPECTROSCOPIA RAMAN AMPLIFICATĂ DE SUPRAFAȚĂ (SERS)

Materia este compusă din molecule cu o distribuție specifică a sarcinii, cum ar fi monopolarile (sarcini izolate) sau dipolii electrici și multipolii superiori (sarcini pozitive și negative). Procesele dintre lumină și materie reprezintă baza oricărei spectroscopii. Câmpul electric al unei unde de lumină poate interacționa cu momentele de dipol existente sau, în cazul moleculelor care nu au moment de dipol permanent, poate induce un moment de dipol prin schimbarea sarcinilor. Într-un fascicul de lumină câmpurile se schimbă de aproximativ 10^{15} ori pe secundă, ceea ce duce la deplasarea (oscilarea) foarte rapidă a moleculelor, rezultatul final fiind generarea de câmpuri electromagnetice. Informațiile obținute din astfel de interacțiuni sunt unice și specifice, deoarece fiecare tip de moleculă are propria distribuție de sarcină, iar schimbul de energie dintre unda electromagnetică și materie este de asemenea specific.

Spectroscopia Raman este o tehnică vibrațională prin care se obțin informații despre structura specifică a moleculelor. Spectroscopia Raman amplificată de suprafață (SERS) este o tehnică Raman care duce la detectarea de semnal Raman mult îmbunătățit de la moleculele active Raman atunci când acestea sunt atașate unor structuri metalice de dimensiuni nanometrice. Factorul de îmbunătățire poate ajunge până la $10^{10} - 10^{11}$, ceea ce înseamnă că se pot detecta chiar și molecule izolate. [1, 2]

1.1. Abordare teoretică

Împăștierea Raman apare doar atunci când variația polarizabilității moleculei este nenulă în timpul vibrației legăturilor moleculare. Polarizabilitatea este o măsură a gradului de deformare a norului electronic molecular sub efectul aplicării câmpului electric incident și depinde de structura moleculei:

$$\mu = \alpha E \quad (1.1)$$

unde α este polarizabilitatea moleculei.

Polarizabilitatea moleculei α sub un câmp electric cu frecvența ν_0 de forma

$$E = E_0 \cos(2\pi\nu_0 t) \quad (1.2)$$

poate fi dezvoltată în serie Taylor:

$$\alpha = \alpha_0 + \left(\frac{\partial\alpha}{\partial q}\right)_0 q + \dots \quad (1.3)$$

iar dacă înlocuim noul α , atunci dipolul indus va fi de forma $\alpha = \alpha_0 + \left(\frac{\partial\alpha}{\partial q}\right)_0 A \cos(2\pi\nu_0 t)$, (1.4)

unde α_0 este polarizabilitatea moleculei în stare de echilibru la poziția de echilibru (R_e), q este deplasarea față de poziția de echilibru $q = R - R_e = A \cos(2\pi\nu_i t)$, iar ν_i este frecvența vibrației moleculei ca urmare a iradierii acesteia cu radiație laser.

În acest caz, momentul de dipol devine

$$\mu = E_0 \alpha_0 \cos(2\pi\nu_i t) + E_0 \left(\frac{\partial\alpha}{\partial q}\right)_0 A \cos(2\pi\nu_i t) \cos(2\pi\nu_0 t) \quad (1.5)$$

$$\mu = E_0 \alpha_0 \cos(2\pi\nu_i t) + \frac{1}{2} E_0 A \left(\frac{\partial\alpha}{\partial q}\right)_0 [\cos 2\pi(\nu_i - \nu_0)t + \cos 2\pi(\nu_i + \nu_0)t] \quad (1.6)$$

Din ecuația (1.6) se observă că lumina emisă oscilează la trei frecvențe diferite. Primul termen indică faptul că lumina emisă are aceeași frecvență ca și fasciculul incident, care corespunde împrăștierii Rayleigh. Însă, în următorii doi termeni, se vede o schimbare între frecvența luminii emise și cea a luminii incidente. Acești termeni corespund împrăștierii Stokes și anti-Stokes. Intensitatea I a luminii emise este proporțională cu E^2 , de aceea împrăștierile Stokes și anti-Stokes se pot exprima:

$$I_{Stokes} = (\nu_i - \nu_0)^4 \cdot \left(\frac{\partial\alpha}{\partial q}\right)_0^2 \cdot I_0(\nu_i) \quad (1.7)$$

$$I_{anti-Stokes} = (\nu_i + \nu_0)^4 \cdot \left(\frac{\partial\alpha}{\partial q}\right)_0^2 \cdot I_0(\nu_i) \cdot e^{\frac{-E_i}{kT}} \quad (1.8)$$

Astfel, fotonii din radiația monocromatică ajunși pe probă pot pierde energie – energie cedată moleculei care este excitată de pe nivelul de bază pe primul nivel vibrațional excitat (împrăștiere Stokes) sau fotonii pot câștiga energie de la molecule care sunt implicate în procesul invers

(împrăștiere anti-Stokes). Dacă energia fascicului incident nu este suficientă pentru a trimite molecula din starea fundamentală în cea mai mică stare electronică, excitația moleculei va avea loc pe o stare virtuală, între cele două stări – Fig. 1.1. Fiind o stare virtuală, molecula nu poate rămâne aici pentru mult timp, așa că va reveni la starea de bază.

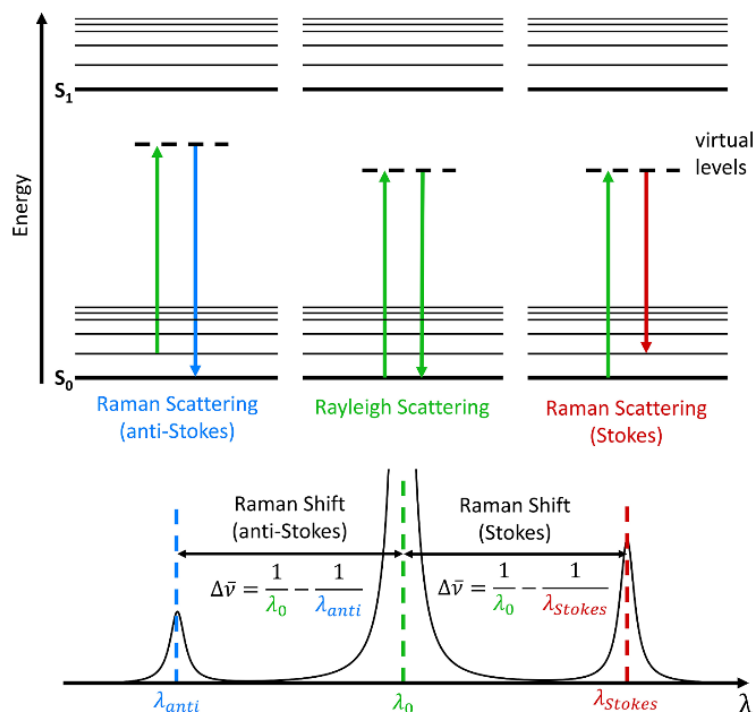


Fig. 1.1 Tranzițiile electronice posibile în spectroscopia Raman: modurile de împrăștiere Rayleigh, Stokes și anti-Stokes

Împrăștierea Raman este, în termeni simpli, alcătuită din ciocniri elastice și inelastice între fotonii incidenti și molecule. Prin ciocnirea elastică, fotonii emiși vor avea aceeași energie $h\nu_i$, la fel ca și cei incidenti, deci frecvența luminii emise va fi aceeași cu cea a luminii incidente, iar molecula se va întoarce în starea din care a plecat. Acesta este procesul cel mai posibil și se numește împrăștierea Rayleigh, dar acești fotoni emiși nu aduc niciun fel de informații legate de moleculele din probă.

Informațiile de orice fel se pot obține prin ciocniri inelastice, deoarece fotonii incidenti ($h\nu_i$) câștigă sau pierde energie, de forma:

$$h\nu_f = h\nu_i + h\nu_0 \quad (1.9)$$

$$h\nu_f = h\nu_i - h\nu_0 \quad (1.10)$$

Așadar, molecula va reveni într-o altă stare vibrațională, diferită de cea fundamentală. Dacă energia emisă este mai mică decât energia incidentă, atunci este cazul împrăștierii Stokes, iar dacă energia emisă este mai mare decât energia incidentă, este cazul împrăștierii anti-Stokes.

Regulile de selecție pentru spectroscopia Raman sunt $\Delta v = -1, 0, +1$, pentru împrăștieria anti-Stokes, Rayleigh, respectiv Stokes, unde Δv este diferența dintre nivelul vibrațional fundamental pe care se află molecula și nivelul vibrațional pe care se întoarce în urma excitării. Pentru ca împrăștieria anti-Stokes să aibă loc, moleculele trebuie să se afle într-o stare de excitație vibrațională (S_1), iar pentru cea Stokes, într-o stare vibrațională de bază (S_0). Fig. 1.2 expune o reprezentare generală a posibilelor căi de transmisie de pe diferite nivele energetice.

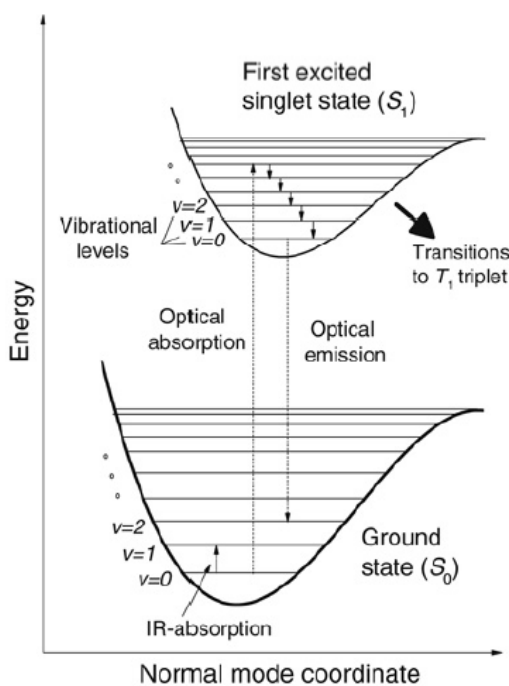


Fig. 1.2 Reprezentarea nivelelor energetice vibraționale ale unei molecule

Doar 1 din 10^6 fotoni sunt împrăștiați inelastic pe probă, sensibilitatea tehnicii Raman fiind scăzută. De aceea, pentru detecția unor concentrații mici de molecule este nevoie de o metodă de

amplificare a semnalului Raman precum: Surface-enhanced Raman scattering (SERS), Tip-enhanced Raman scattering (TERS), Surface plasmon polariton enhanced Raman scattering (SPPERS), Surface-enhanced resonance Raman spectroscopy (SERRS).

Efectul SERS a fost observat pentru prima dată în 1974, de către Fleischman et al. [19], datorită apariției unui semnal Raman neobișnuit de ridicat provenit de la piridina plasată pe un electrod dur de argint. Fenomenul a fost explicat printr-o creștere în suprafață a electrodului dur, ceea ce înseamnă o creștere numerică a moleculelor adsorbite, respectiv o îmbunătățire a semnalului Raman. În 1977, Jeamarie și Van Duyne și, independent, Albrecht și Creighton, au confirmat experimentele din 1974, dar au descoperit că acest semnal Raman îmbunătățit este cauzat de o intensificare a secțiunii efective Raman și că nu poate fi explicat doar pe baza creșterii numerice a moleculelor adsorbite.

În perioada următoare, s-au analizat astfel de semnale Raman îmbunătățite pentru diferite molecule adsorbite de suprafețe rugoase de argint, aur sau cupru, iar tehnica s-a numit Spectroscopia Raman amplificată de suprafață – *surface-enhanced Raman spectroscopy* sau *surface-enhanced Raman scattering* (SERS). În primele experimente, factorul de îmbunătățire era estimat la un ordin de $10^3 - 10^5$, în prezent ajungând la $10^{10} - 10^{11}$. [1, 2]

Într-un experiment SERS moleculele sunt adsorbite de nanostructuri metalice. Cele mai frecvente nanostructuri care oferă cele mai mari efecte de îmbunătățire a semnalului sunt particulele de coloid de argint sau aur cu dimensiunea de 10 – 100 nm, electrozi cu suprafețe rugoase de aur sau argint sau pelicule insulare evaporate ale acestor metale și sunt numite *substraturi active SERS*.

Există două teorii diferite, susținute și argumentate, ce ar putea sta la baza efectului de amplificare Raman, însă niciuna nu explică întreg ansamblul: teoria electromagnetică și teoria chimică. Ambele teorii confirmă totuși necesitatea unor structuri metalice cu mărimi nanometrice în apropierea moleculei, sau chiar pe care molecula să adsoarbă.

Teoria electromagnetică este o caracteristică a substratului și nu ține cont de molecula țintă. Apare ca urmare a localizării luminii la suprafața substratului, mai exact a cuplării fotonilor cu electronii de suprafață. Vibrația colectivă a acestor electroni din păturile superioare ale nanostructurii metalice este numită plasmon. Dacă radiația laser incidentă îndeplinește condiția de rezonanță plasmonică, va fi generat un câmp electromagnetic de către mișcarea colectivă a plasmonilor de suprafață – Fig. 1.3.

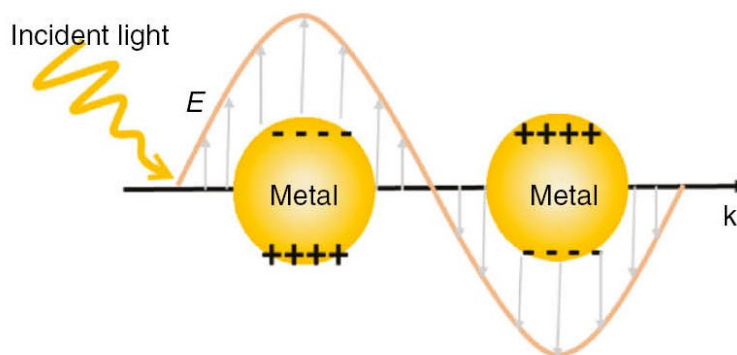


Figura 1.3 Rezonanța plasmonilor de suprafață

Molecula trebuie să se afle în apropierea nanostructurii, mai exact în zona de acțiune a câmpului electromagnetic generat de vibrația plasmonică (1-10 nm), iar câmpul electromagnetic astfel generat va modifica eficiența împrăștierii Raman deoarece puterea radiației împrăștiate de un dipol este dependentă de mediul în care acesta se află.

Teoria chimică în schimb necesită o interacțiune fizico-chimică între moleculă și substratul metalic, depinzând astfel de tipul moleculei și nivelele ei energetice. Distanța dintre moleculă și substratul metalic în acest caz este de ordinul a câțiva Angstromi.

Mecanismul chimic implică la rândul lui două tipuri de interacțiuni între moleculă și substrat: fizisorpția și chemisorpția. Fizisorpția reprezintă o interacțiune slabă de tip van der Waals între moleculă și substratul metalic, fără schimbarea nivelelor energetice ale moleculei. Chemisorpția însă constă în formarea unui ansamblu moleculă-metal, prin hibridizarea metalului și orbitalilor moleculari ai absorbantului pentru a forma stări de legătură și anti-legătură. Efectul SERS constă în transferul electronic rezonant între nivelele HOMO (highest occupied molecular orbitals) și LUMO (lowest unoccupied molecular orbital) al complexului metal-moleculă. Astfel, efectul chimic poate explica amplificarea preferențială a modurilor de vibrație ale moleculelor adsorbite pe suprafața metalică.

Ambele teorii, și cea electromagnetică, și cea chimică, au fost dovedite și sunt destul de greu de diferențiat. Cert este că mecanismul chimic ar trebui să apară numai datorită moleculelor direct atașate la suprafață și, prin urmare, ar trebui să crească numai până la acoperirea primului strat, iar cel electromagnetic, deși are un efect mai lung, scade cu $1 / r^3$ o dată cu distanțarea față de suprafață. Astfel, majoritatea îmbunătățirilor pe bază electromagnetică vor apărea, de asemenea, tot prin prezența primului strat la suprafața metalului.

1.2. SERS în medicină

Tehnica SERS a fost introdusă în ramura biomedicinii încă din anii 80-90 odată cu detecția unei singure molecule de hemoglobină [4] și a detecției acizilor nucleici [20] folosind spectroscopia SERS. Datorită limitei joase de detecție, vitezei de aplicare, capacității de multiplexare și a sensibilității ridicate a tehnicii, în ultimul deceniu a existat o creștere rapidă a numărului de aplicații de spectroscopie SERS în cercetarea medicală, în special în diagnosticul medical.

Biopsia lichidă SERS este o abordare în care algoritmi de machine learning (ML) permit clasificarea probelor în pacienți și control pe baza semnalului SERS al biofluidelor și a demonstrat progrese semnificative în ultimul deceniu [3-5]. S-a demonstrat că SERS ar putea completa tehnici multiple de diagnostic, cum ar fi histopatologia pentru diagnosticul cancerului [4, 21], rezonanța magnetică nucleară pentru diagnosticul leziunilor cerebrale traumatiche [22], biopsia hepatică pentru bolile hepatice [23] sau reacția în lanț a polimerazei pentru depistarea precoce a conjunctivitei adenovirale [24]. Dintre diferitele abordări SERS pentru diagnosticul medical, biopsia lichidă SERS a fost până acum cea mai populară alegere, deoarece este minim invazivă, necesită volume mici de probă și este o tehnică rapidă de achiziție.

Precizia generală a diagnosticului medical bazat pe SERS atinge la momentul actual o valoare de ~80–90% [7, 15-18]. Deși această valoare rămâne sub unele tehnici standard de laborator clinic, ea evidențiază potențialul SERS ca instrument versatil în cadrul point-of-care, fiind o metodă de screening rapidă și ușor de utilizat.

Biopsia lichidă SERS s-a dovedit utilă în diagnosticarea a numeroase boli, arătând că există trei clase principale de molecule care pot fi detectate de SERS: metaboliți purinici, proteine și carotenoizi care furnizează informații cruciale cu privire la rata de refacere celulară, inflamație și, respectiv, capacitatea antioxidantă, procese ce sunt perturbate în aproape fiecare boală – Fig. 1.4. Studii recente au arătat că, la spectrele SERS ale serului filtrat și deproteinizat sau plasmei mai contribuie și un aminoacid alimentar, ergotioneina [25].

În ceea ce privește detectarea cancerului prin spectroscopie SERS, a fost raportată o acuratețe de peste 85% în diagnosticarea cancerului din probe biologice lichide. Detectarea anomaliilor maligne se realizează prin analiza ARN-ului sau ADN-ului celular sau prin urmărirea modificărilor

concentrației metaboliților purinici (xantină, hipoxantină, acid uric) care sunt secretați în cantități ridicate în contextul malignității. Există studii în care discriminarea s-a realizat pe baza modificărilor de concentrație și, implicit, a semnalului SERS a metaboliților purinici din urină între pacienți cu cancer de sân și pacienți control cu o acuratețe de 88% [8]. Când probele de ser sangvin filtrat au fost analizate prin SERS, acuratețea în distingerea celor două categorii de probe a fost mai mare (93,6%).

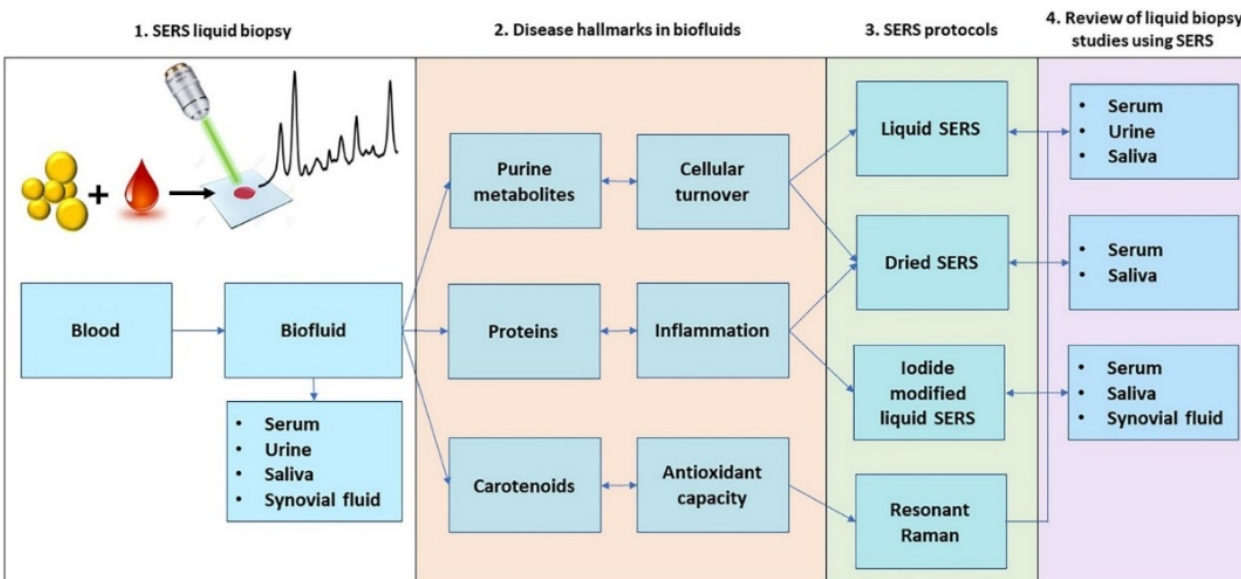


Fig. 1.4 - Relația dintre clasa de analiți, rolul lor funcțional și protocoalele experimentale SERS utilizate pentru detectarea lor [26]

Prin achiziționarea spectrelor SERS ale serului sangvin și analiza lor prin tehnica de clasificare PCA-LDA, s-a atins o acuratețe totală de 73.7% în cazul cancerului colorectal, 85.75% pentru cancerul de plămâni, 95% în cazul cancerului ovarian, 92% pentru cancerul oral [27] și 90% pentru cancerul pancreatic [28]. Acuratețe ridicată a fost raportată și în cazul analizării SERS a probelor de țesut extrase din tumoare, pentru cancerul de vezică acuratețea atingând o valoare de 94% [29] iar în cazul cancerului oral 97.24% [30].

Pentru a îmbunătăți diagnosticul medical minim invaziv pe baza spectroscopiei SERS, datele spectrale pot fi combinate cu valorile unor parametri medicali detectați în sânge sau urină. De exemplu, prin combinarea valorilor antigenului specific de prostată (PSA) cu spectrele SERS ale serului sangvin filtrat, acuratețea diagnosticului cancerului de prostată bazat pe PSA a fost îmbunătățită de la 70.4%

la 94%. [21] Prin introducerea antigenului carcinoembrionar ca factor discriminant, acuratețea detectării cancerului colorectal pe baza spectrelor SERS ale serului sangvin a fost îmbunătățită de la 89.8% la 91.8% [31].

Implicațiile SERS în evoluția medicinei sunt enorme, prin urmare alegerea corectă a substratelor metalice în vederea obținerii unui semnal SERS este foarte importantă. Detecția moleculelor biologice poate fi realizată fie indirect, folosind substraturi metalice funcționalizate care să formeze legături cu molecula dorită, fie chiar direct, molecula fiind adsorbită pe suprafața metalică prin grupări cu afinitate mare față de metalul folosit sau prin adăugarea unor săruri anorganice (ex. KI, MgSO₄, NaCl, CaNO₃ etc.) care să faciliteze această interacțiune metal-analit [27, 32]. Nanoparticulele metalice sunt cele mai răspândite substraturi metalice utilizate în SERS, iar cele mai multe aplicații biomedicale ale SERS au folosit ca substrat metalic nanoparticule de argint (AgNPs).

În SERS, numai moleculele adsorbite pe nanoparticulele metalice plasmonice pot genera semnal. Prin urmare, adsorbția controlată a analiților pe suprafețele metalice este crucială pentru obținerea unor rezultate fiabile. Studiile recente [2, 33-35] propun un model de adsorbție specific ionilor adsorbiți, care arată că ad-ionii (ionii adsorbiți) permit controlul adsorbției analiților cationici și anionici în moduri specifice pe nanoparticule de argint: adioni cationici (Ca²⁺, Mg²⁺, Pb²⁺ sau Al³⁺) promovează chimisorbția moleculelor anionice, în timp ce ad-ionii anionici (Cl⁻, Br⁻ sau I⁻) promovează chimisorbția moleculelor cationice pe nanoparticulele de argint. Pe lângă adsorbția specifică, efectul selectiv de *activare* SERS al speciilor anionice, are loc într-o manieră competitivă, în funcție de afinitatea acestor specii pentru suprafața metalului – Fig. 1.5.

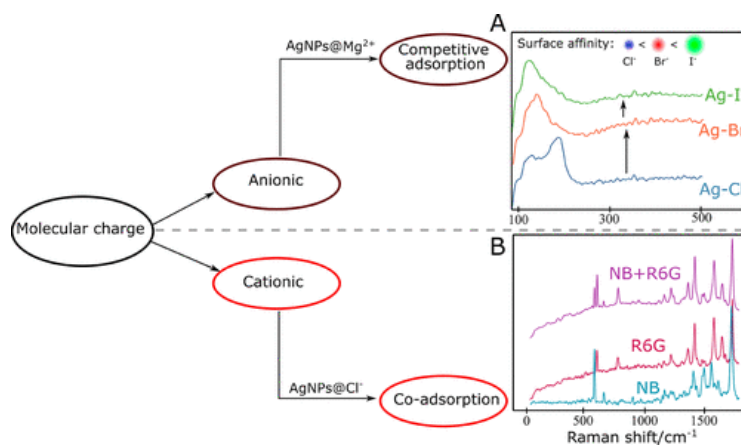


Fig. 1.5 Mecanismul specific de adsorbție al analitului țintă la AgNP [2]

Spectrele SERS ale biofluidelor umane nu au atribuiri complete ale benzilor vibraționale. În plus, pot apărea modificări spectrale din cauza neomogenității substratului metalic utilizat sau a diferitelor tipuri de legături dintre molecule și substratul metalic. Pentru ca aceste fluctuații să nu interfereze cu diagnosticul medical, s-au dezvoltat diverse metode de analiză statistică folosite în analiza spectrelor SERS elimină erorile sistematice.

CAPITOLUL 2.

CANCERUL DE SÂN (MAMAR)

Cancerul de sân afectează 1 din 7 (14%) dintre femeile din întreaga lume. De asemenea, la femei, cancerul de sân reprezintă 22.9% dintre cazurile de cancer invaziv și 16% dintre toate cazurile de cancer. Numărul de cazuri a crescut semnificativ începând cu anii '70, fenomen atribuit în parte stilurilor de viață moderne. În 2011 au fost înregistrate mai mult de 41.000 de cazuri recent diagnosticate în Anglia. Pe baza statisticilor americane, în 2015 existau 2.8 milioane de femei afectate de cancer de sân. În 2018 a rezultat în 2 milioane de cazuri noi și 627.000 de decese. [36]

La fel ca orice tip de cancer, și cancerul de sân apare din cauza unei interacțiuni dintre un factor de mediu (extern) și o gazdă susceptibilă genetic. Celulele normale se divid de câte ori este necesar și se opresc, se atașează de alte celule și rămân pe loc în țesuturi. Ele devin canceroase atunci când își pierd aceste capacități.


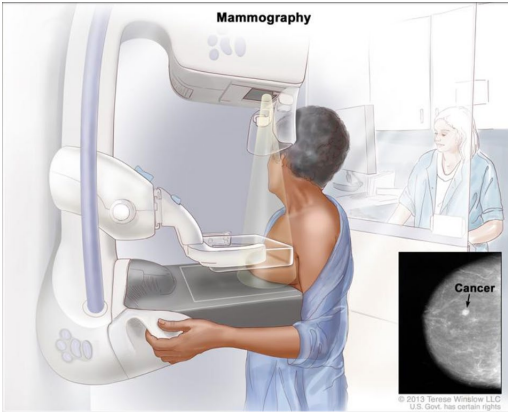
Semnele de cancer de sân cele mai frecvente constau în apariția unui nodul la nivelul sânului, adică o porțiune de țesut mamar se simte diferit de restul. Peste 80% din cazuri sunt descoperite atunci când o persoană detectează tactil o astfel de modificare. Cu toate acestea, pentru o detectare timpurie, este necesară realizarea unei mamografii. Bulele găsite în ganglionii limfatici localizați în axile pot indica, de asemenea, cancer de sân.

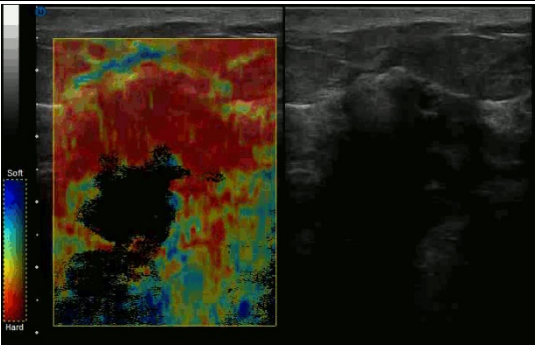
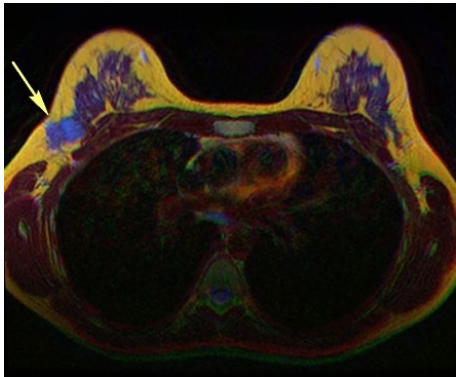

Un alt simptom al acestui tip de cancer este boala Paget a sânului, prin care se observă modificări ale pielii asemănătoare cu eczema (roșeață, decolorare). Pe măsură ce boala Paget avansează, simptomele pot include furnicături, mâncărime, sensibilitate crescută, arsură și durere. Aproximativ jumătate dintre femeile diagnosticate cu boala Paget a sânului au, de asemenea, un nodul la sân. [37]

Factorii de risc pentru dezvoltarea cancerului de sân pot fi împărțiți în două categorii: factori de risc modificabili (obezitatea, lipsa exercițiilor fizice, alcoolismul) și ficși, precum sexul biologic (femeile sunt mult mai predispuse acestui tip de cancer, dar poate fi întâlnit și în rândul bărbaților), terapia de înlocuire hormonală în timpul menopauzei, apariția târzie (sau deloc) a copiilor, un istoric anterior al cancerului de sân și un istoric familial de cancer de sân. Aproximativ 5–10% din cazuri sunt

rezultatul unei predispoziții genetice moștenite de la părinții unei persoane, incluzând BRCA1 și BRCA2 (mutații specifice cancerului de sân). [38] Cele mai frecvent utilizate metode de screening ale cancerului de sân sunt reprezentate în Tabelul 2.1.

Tabelul 2.1 Cele mai frecvent utilizate metode de screening ale cancerului de sân

<ul style="list-style-type: none"> • examinarea clinică prin vizualizare și palpare; 	
<ul style="list-style-type: none"> • mamografia (radiografie a sânului) – sânul este presat între două plăci, iar razele X trec prin țesut nedeiate; dacă se descoperă o zonă densă, în care nu se întâmplă acest lucru, atunci acolo se presupune a exista o tumoră; • o revizuire Cochrane din 2013 a constatat că nu este clar dacă screeningul mamografic face mai mult rău decât bine, în condițiile în care o mare parte dintre femeile care au fost testate pozitiv s-au dovedit a nu avea boala; [39] 	

<ul style="list-style-type: none"> • examinarea cu ultrasunete: undele sonore cu energie mare (ultrasunete) sunt respinse de țesuturile sau organele interne și fac ecouri; 	
<ul style="list-style-type: none"> • studii asupra compoziției chimice a sângelui: este verificată o probă de sânge pentru a măsura cantitățile din anumite substanțe eliberate în sânge de organe și țesuturi; 	
<ul style="list-style-type: none"> • examinarea RMN: cu ajutorul unui câmp magnetic extern și a radiofrecvențelor, se obțin imagini ale structurii interne a corpului uman; 	
<ul style="list-style-type: none"> • biopsie: îndepărtarea celulelor (în special din nodul dacă se depistează) astfel încât acestea să poată fi analizate la microscop; • există mai multe tipuri de biopsii: excizională (îndepărtarea unei întregi bucăți de țesut), incizională (îndepărtarea unei părți dintr-un lot sau a unui eșantion de țesut), de bază (îndepărtarea țesutului 	

folosind un ac larg) sau de aspirație cu ac fin – FNA (îndepărtarea țesutului sau a fluidului, folosind un ac subțire).	
---	--

În urma testelor se obțin informații precum cât de repede poate crește tumora, cât de probabil este să se răspândească în corp (stadiul în care se află), ce metode de tratament ar funcționa cel mai bine, etc. Astfel, se ajunge la diagnosticul final și la selectarea celor mai eficiente metode de tratament (chirurgie, radioterapie, chimioterapie, terapie hormonală și terapie vizată – cu anticorpi).

Screeningul eficient și precoce urmat de o intervenție chirurgicală și chimioterapie sau radioterapie contribuie la o scădere semnificativă a mortalității. În prezent, cea mai frecvent utilizată metodă de screening este examinarea mamografică. Deși este foarte eficientă pentru reducerea ratei de deces, mamografia suferă în continuare de unele limitări, cum ar fi sensibilitate scăzută în țesutul mamar cu densitate mai mare, alarmă fals pozitivă frecventă și risc de iradiere. [39]

Prin urmare, există o nevoie semnificativă de noi tehnici în screening-ul medical care să aibă un nivel ridicat de accesibilitate, sensibilitate, specificitate și acuratețe. Tehnica SERS este analiza spectroscopică care se încadrează în criteriile menționate anterior și oferă informații despre compoziția metabolică a biofluidelor. Combinată cu algoritmi de machine learning (ML) în vederea discriminării probelor, SERS poate deveni un instrument puternic pentru caracterizarea sistemele biologice complexe cu scopuri diagnostice și terapeutice.

CAPITOLUL 3.

MATERIALE ȘI METODE

3.1. Colectarea, pregătirea și depozitarea probelor de ser și urină

Studiul prezent include 60 de probe de ser și urină obținute de la aceleași pacienți, dintre care 39 au fost colectate de la pacienți cu cancer de sân (confirmat prin biopsie) și 21 de probe de la pacienți control. De la fiecare pacientă, serul și urina au fost colectate dimineața și păstrate la $-80\text{ }^{\circ}\text{C}$ până la analiză.

În vederea obținerii spectrelor SERS ale metaboliților din ser, probele de biofluid au fost filtrate și deproteinizate, pentru a preveni adsorbția proteinelor pe suprafețele metalice și astfel formarea unor coroane proteice în jurul nanoparticulelor de argint (AgNP) care ar bloca adsorbția metaboliților de interes [40]. Astfel, probele de ser au fost deproteinizate prin precipitare cu metanol. Pentru aceasta s-au amestecat $900\text{ }\mu\text{L}$ metanol și $100\text{ }\mu\text{L}$ biofluid, după care au fost centrifugate 15 minute la $5800\times g$ (unități gravitaționale) pentru a separa precipitatul proteic de supernatantul care conține metaboliții. Prin urmare, probele analizate au fost reprezentante de fracțiunea metabolică a serului, fără conținut de proteine. Supernatantul a fost colectat și păstrat la $-20\text{ }^{\circ}\text{C}$ până la măsurare.

3.2. Substratul SERS

Spectrele SERS au fost obținute folosind ca și substrat nanoparticule de argint reduse cu clorhidrat de hidroxilamină preparate prin metoda Leopold-Lendl, avantajul acestuia fiind prepararea rapidă la temperatura camerei și aplicabilitatea imediată. [41]

Metoda Leopold-Lendl presupune dizolvarea a 0.017 g de azotat de argint în 90 ml apă Milli-Q (apă ultrapură). În paralel, se prepară o soluție de 10 ml apă Milli-Q și 1.2 ml hidroxid de sodiu (concentrație 1%) în care s-a dizolvat 0.02 g clorhidrat de hidroxilamină. Cele două soluții, azotatul de argint și compusul cu hidroxilamină, se amestecă sub agitare. Instantaneu, soluția finală (coloidul) devine galben închis. Această culoare marchează faptul că ionii de argint au fost reduși de

hidroxilamină și s-au format nanoparticule. Soluția de hya-AgNPs a fost păstrată la temperatura camerei.

Spectrele UV-Vis ale coloidului de argint au fost înregistrate folosind un spectrofotometru Jasco V-630 UV-Vis, cu ajutorul unor cuve de cuarț cu o cale optică de 1 cm (Fig. 3.1 A.) și oferă o indicație asupra concentrației și distribuției mărimii nanoparticulelor. [42] Dimensiunea mai precisă a nanoparticulelor de argint (30-50 nm) a fost determinată folosind microscopia electronică de transmisie (TEM) – Fig. 3.1 B. De asemenea, stabilitatea nanoparticulelor a fost testată prin determinarea potențialului Zeta al nanoparticulelor, care atinge o valoare de -31.7 ± 9.29 mV – Fig. 3.3 C.

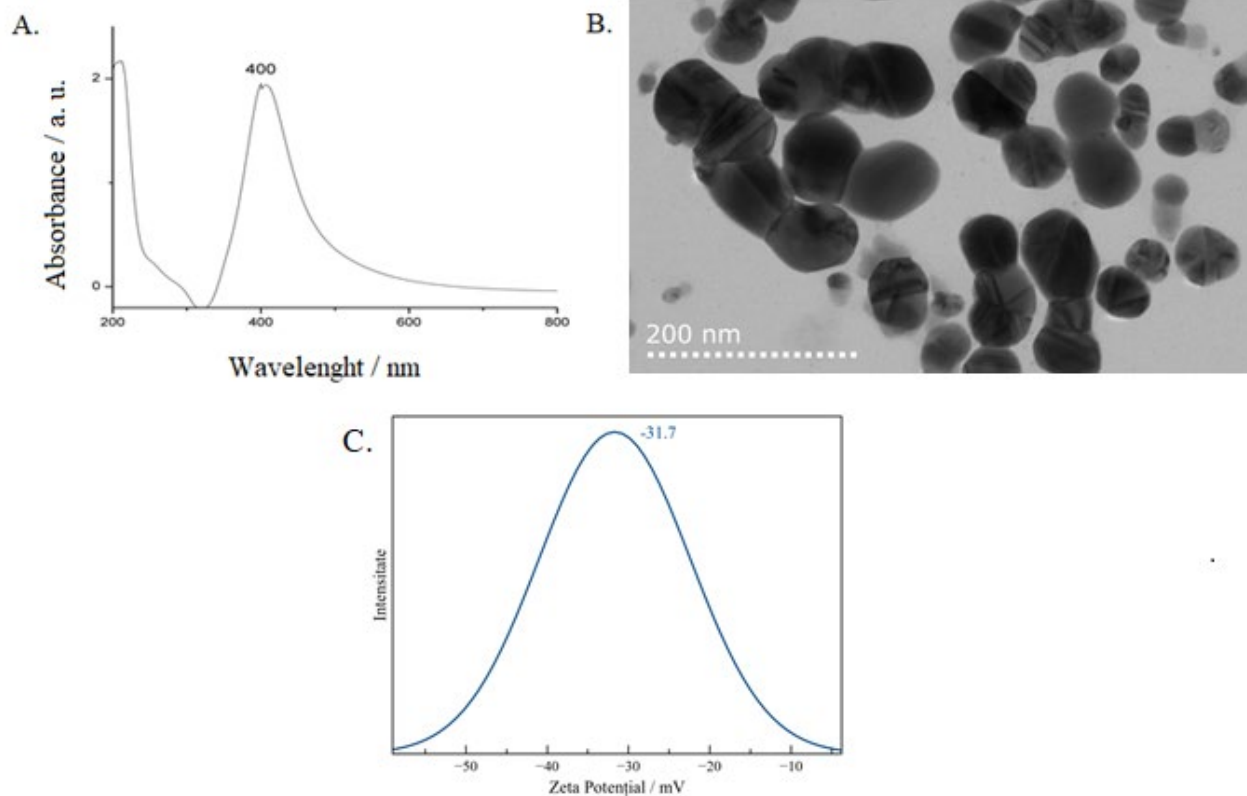


Fig. 3.1 A. Spectrul UV-Vis al nanoparticulelor hya-AgNPs B. Imaginea TEM a nanoparticulelor hya-AgNPs C. Potențialul Zeta al nanoparticulelor hya-AgNPs

3.3. Metodologia SERS utilizată

Metodologia măsurătorilor SERS pentru detecția analiților metabolici din serul sangvin filtrat și deproteinizat a devenit în ultimul timp un subiect de interes, publicațiile recente susținând importanța ad-ionilor în detecția SERS a moleculelor anionice și cationice [33-35]. Înregistrarea spectrelor SERS ale acizilor organici (ex. acidul uric, acidul salicilic) este mediată de adsorbția cationilor precum Ca^{2+} , Al^{3+} sau Pb^{2+} pe anumite defecte de suprafață ale nanoparticulelor de argint (AgNPs), crescând astfel locurile de adsorbție disponibile pentru analiții anionici [33]. Prin urmare, detecția SERS metaboliților purinici, analiți anionici, din serul sangvin filtrat necesită chemisorpția lor pe suprafața nanoparticulelor de argint prin cationi.

Pentru a favoriza adsorbția metaboliților anionici din urină sau ser filtrat pe AgNPs, am amestecat, prin pipetare, 90 μL hya-AgNPs (coloid), 10 μL soluție de biofluid și 10 μL $\text{Ca}(\text{NO}_3)_2$ 10^{-2} M (concentrație de Ca^{2+} din amestecul final a fost de 10^{-4} M). Pentru concentrații de Ca^{2+} mai mici de 0.2 mM soluția coloidală a rămas stabilă și nu a fost observată nicio agregare.

Pentru achiziția spectrelor am folosit o linie laser de 532 nm (putere laser de 20 mW pe probă), fiind focalizat pe probe printr-un obiectiv de microscop 5X (NA=0.12. Pentru fiecare probă s-au înregistrat câte 3 spectre într-un timp de achiziție de 40 secunde fiecare, mediate apoi într-un singur spectru final. O picătură de 5 μL din probă a fost plasată pe o lamă de microscop acoperită cu folie de aluminiu, urmând a fi analizată conform descrierilor de mai sus. Am folosit folie de aluminiu pentru a evita orice fel de interferențe spectrale ce ar putea apărea din cauza sticlei. Măsurătorile au fost realizate într-un mediu clinic, iar spectrele au fost înregistrate aleatoriu.

3.4. Pre-procesarea și analizarea spectrelor SERS

Pre-procesarea și analizarea datelor au fost efectuate folosind softul Quasar – Fig. 3.2 [43]. Pre-procesarea spectrelor presupune tăierea spectrelor și selectarea intervalului de interes, 300 – 1800 cm^{-1} , urmată de eliminarea zgomotului de fundal prin corecția liniară de bază (*linear baseline correction*). Următorul pas a fost normalizarea spectrelor, folosind normalizarea medie (*mean normalization*) – împărțirea valorilor la valoarea lor mediană. După aceea a urmat uniformizarea

(*smoothing*) cu ajutorul algoritmului Savitzky-Golay. Acesta atașează un polinom la fiecare segment de curbe succesive, înlocuind valorile originale cu variații mai regulate. Astfel, se utilizează pentru a elimina eficient peak-urile de zgomot spectral în timp ce informațiile chimice sunt păstrate.

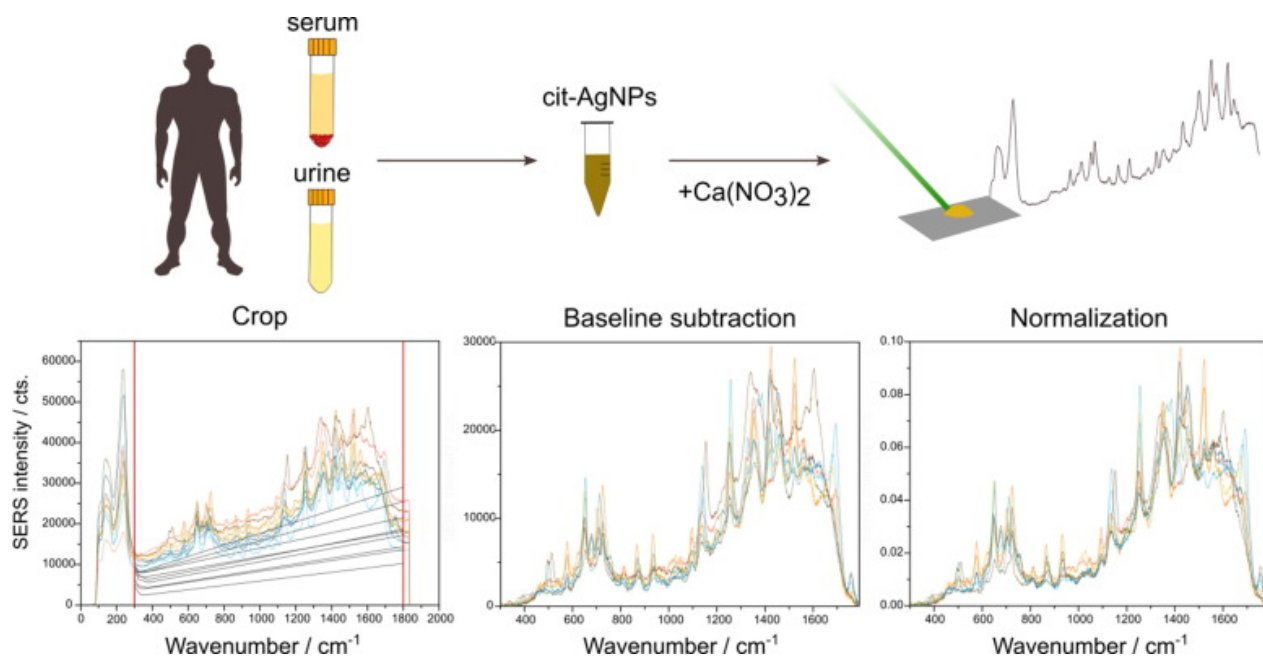


Fig. 3.2 Workflow-ul utilizat pentru analiza SERS a biofluidelor ser sangvin și urină

PCA (*Principal Component Analysis*) este un tip de analiză nesupervizată și a fost utilizată pentru a examina diferențele dintre spectrele SERS ale fracțiunii metabolice din ser filtrat și urinei. PCA este o metodă de analiză statistică prin care se poate reduce dimensionalitatea spectrelor prin calcarea unor noi variabile ortogonale numite *componente principale* (PC), păstrând în același timp cât mai multe informații posibil. Fiecare PC reprezintă o combinație liniară a variabilelor inițiale (în acest caz, numerele de undă ale spectrului SERS), astfel încât informația redundantă să fie eliminată. Matematic, PCA este definită ca o transformare liniară ortogonală prin care datele dintr-un sistem se transferă într-un nou sistem de coordonate, astfel încât cea mai mare varianță de pe o anumită proiecție a datelor se află pe prima coordonată (numită prima componentă principală), a doua cea mai mare varianță este localizată pe a doua componentă principală și așa mai departe.

Tabelele mari de date conțin o cantitate mare de informații, care este uneori parțial ascunsă, deoarece datele sunt prea complexe pentru a putea fi interpretate cu ușurință. Astfel, PCA este o metodă

de proiecție care ajută la vizualizarea tuturor informațiilor conținute într-un set de date, oferind rezultate precum: în ce privință un eșantion este diferit de altul, ce variabile contribuie cel mai mult la această diferență, dacă acestea contribuie în același mod (adică sunt corelate) sau independent una de cealaltă sau detectarea unor tipare de eșantion (orice grupare particulară), etc. Cu alte cuvinte, evidențiază informațiile utile conținute în date.

3.5. Algoritmii de machine learning (ML)

În acest studiu, am explorat cinci metode de clasificare supravegheată, care presupun pre-etichetarea probelor. Informațiile sunt împărțite în două seturi, care sunt utilizate în continuare pentru antrenarea și validarea clasificatorilor. La selectarea eșantioanelor corespunzătoare seturilor menționate mai sus, trebuie păstrat raportul claselor din datele originale. Pentru acest lucru am folosit metoda *cross validation*, care presupune utilizarea a 80% dintre spectrele colectate ale fiecărui subset (Control – CTRL sau Breast Cancer – BC) pentru a antrena modelul de clasificare, în timp ce spectrele din eșantioanele rămase au fost considerate date noi.

Cele cinci clasificatoare incluse în acest studiu au fost Decision Tree (DT), Random Forest (RF), Gaussian Naïve Bayes (GNB), Support Vector Machine (SVM) și Linear Discriminant Analysis (LDA). DT-urile sunt simple clasificatoare, construind un set de condiții dacă-atunci-altfel pentru a clasifica eșantioanele în clasele specificate. RF, după cum sugerează și numele, conține un număr dat de DT-uri randomizate, iar predicția finală se face pe baza clasei majoritare prezisă de DT-uri. Algoritmii GNB se bazează pe teorema Bayes, presupunând o independență puternică între caracteristici și oferă o metodă de clasificare probabilistică simplă, dar foarte funcțională. SVM-urile se bazează pe cadre de învățare statistică și reprezintă una dintre cele mai robuste metode de clasificare, transformând caracteristicile în puncte într-un spațiu n-dimensional și găsind un hiperplan care ar separa clasele. LDA-urile sunt o generalizare a discriminantului liniar al lui Fisher, care funcționează prin calcularea statisticilor rezumative pentru caracteristicile de intrare pentru a găsi o combinație liniară care separă clasele.

În acest studiu, pentru a evalua performanțele acestor algoritmi, am analizat acuratețea lor generală, precizia, recall-ul și scorul F1 pentru fiecare clasificator în parte. Precizia reprezintă

fracțiunea de elemente relevante dintre elementele recuperate, recall-ul reprezintă fracțiunea elementelor recuperate, în timp ce acuratețea este o măsură statistică prin care se cuantifică proporția de predicții corecte din numărul total de cazuri examinate:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3.1)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3.2)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3.3)$$

unde TP (*true positive*) reprezintă numărul de elemente etichetate corect ca aparținând clasei pozitive, TN (*true negative*) reprezintă numărul de elemente etichetate corect ca aparținând clasei negative, FP (*false positive*) – numărul elementelor care aparțin clasei negative, dar au fost etichetate incorect ca aparținând clasei pozitive, iar FN (*false negative*) – numărul elementelor care aparțin clasei pozitive, dar au fost etichetate incorect ca aparținând clasei negative – Fig. 3.3.

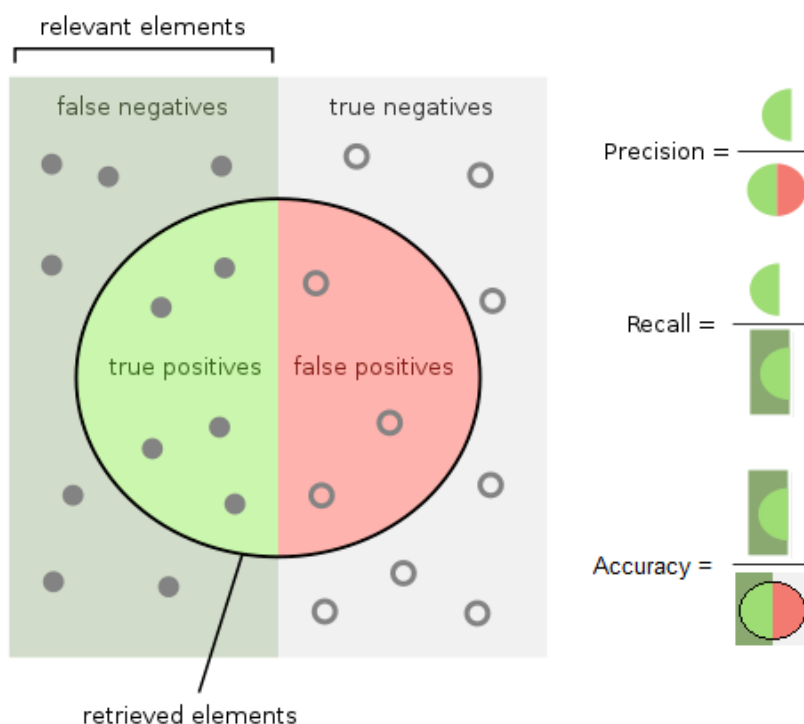


Fig. 3.3 – Precizia, recall-ul și acuratețea în analiza statistică

În analiza statistică, acuratețea testelor mai poate fi măsurată prin *F-measure* sau *F-score*. Scorul F_1 reprezintă media armonică dintre precizie și recall. Cea mai mare valoare posibilă a unui scor F_1 este 1.0, indicând o precizie și un recall perfecți, iar cea mai mică valoare posibilă este 0.0, caz în care fie precizia, fie recallul, fie ambele sunt zero. Scorul F_1 este cunoscut și sub denumirea de coeficient Sørensen–Dice sau coeficient de similitudine Dice (DSC):

$$F_1 = \frac{2}{\text{recall}^{-1} + \text{precision}^{-1}} = 2 \frac{\text{recall} \cdot \text{precision}}{\text{recall} + \text{precision}} = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \quad (3.4)$$

CAPITOLUL 4.

REZULTATE ȘI DISCUȚII

Studiul de față urmărește să atingă două scopuri, redate în următoarele două subcapitole:

1. analiza comparativă și complementară a informațiilor metabolice furnizate de spectrele SERS ale probelor de ser sangvin filtrat și urină de la pacienți diagnosticați cu cancer de sân și pacienți control;
2. analiza acurateții relative a 5 algoritmi de machine learning (ML) aplicați setului de date pentru clasificarea probelor în canceros vs necanceros.

4.1. Analiza comparativă și complementară a informațiilor metabolice furnizate de spectrele SERS ale fracțiunii metabolice din ser filtrat și urinei

În această secțiune, am pus accentul pe analiza calitativă a spectrelor SERS ale fracțiunii metabolice din ser filtrat și urinei, indiferent de apartenența la grup (cancer de sân sau control) pentru a determina și compara informațiile biochimice ale acestora.

Spectrele medii SERS ale fracțiunii metabolice din ser filtrat și urinei din toate probele, indiferent de grupul lor (cancer de sân sau control), împreună cu spectrul lor de diferență, care evidențiază distincțiile spectrale SERS relevante dintre ser și urină, sunt prezentate în Fig. 4.1.

Zona normalizată de sub spectru a fost utilizată ca măsură a variabilității spectrelor SERS din probele de ser filtrat și urină (Fig. 4.1 sus-stânga). Spectrele SERS ale urinei au prezentat o variabilitate puțin mai mare (cu o mediană de 38.72, în comparație cu spectrele fracțiunii metabolice ale serului – 38.50), posibil din cauza intervalului larg de pH din urină comparativ cu cel din ser.

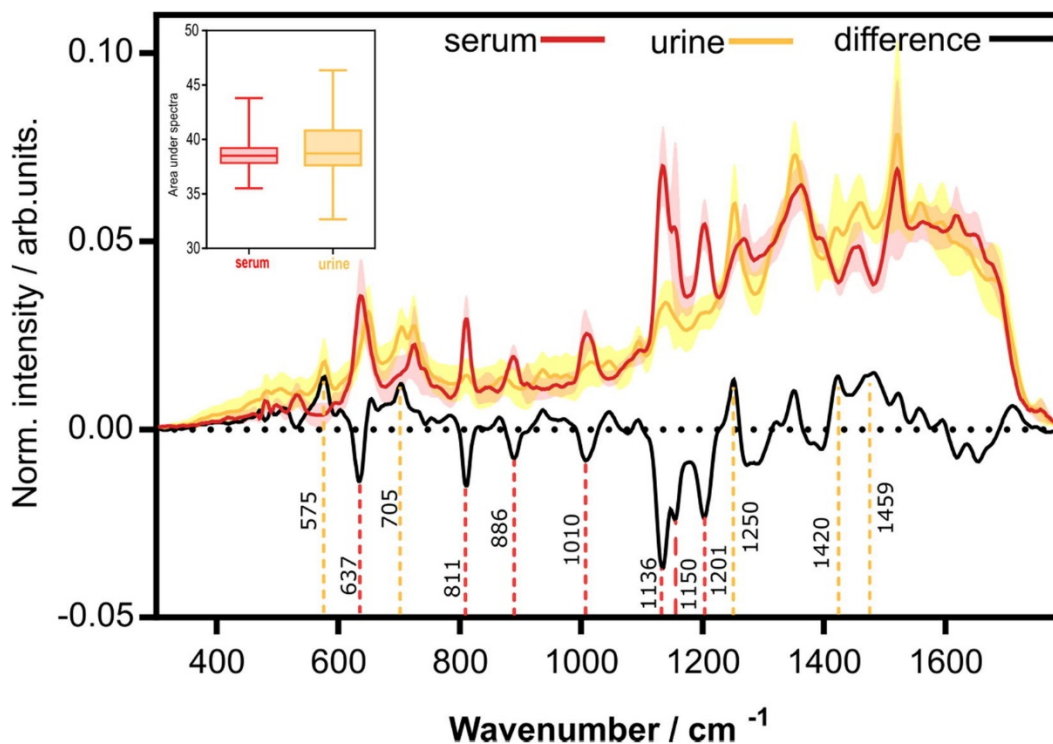


Fig. 4.1 - Spectrele medii SERS ale serului filtrat (roșu) și urinei (galben) împreună cu diferența dintre ele (negru)

O atribuire a benzilor SERS din spectrele probelor de ser filtrat și urină este prezentată în Tabelul 4.1. Atribuirile au fost făcute pe baza studiilor din literatură, precum și pe baza spectrelor SERS ale acidului uric, xantinei și hipoxantinei. Cele mai intense benzi SERS prezente atât în probele de ser filtrat, cât și de urină, sunt observate la 637 cm^{-1} , 811 cm^{-1} , 889 cm^{-1} , 1136 cm^{-1} și 1201 cm^{-1} , care pot fi toate atribuite acidului uric, așa cum susțin și Bonifacio și colab.[40], Westley și colab. [44] și Moisoiu și colab. [9]. Hipoxantina, un precursor al acidului uric, se observă atât în spectrul SERS seric, cât și în urină prin vârful la 725 cm^{-1} , în timp ce pentru xantină, forma redusă a hipoxantinei, atribuim umărul la 1250 cm^{-1} , care se observă doar în probe de urină. Benzile SERS la 1368 cm^{-1} și 1686 cm^{-1} nu au putut fi atribuite unui singur metabolit, deoarece acestea sunt observate în fiecare dintre spectrele SERS ale acidului uric, xantinei și hipoxantinei.

Tabel 4.1 Concentrația normală a metaboliților în ser și urină, respectiv atribuirea benzilor SERS

Metabolite	Serum level (μM)	Urine level ($\mu\text{mol}/\text{mmol}$ creatinine)	SERS bands assignment (cm^{-1})		
			Serum	Urine	Serum and urine
Ergothioneine	0.84-1.52		482		499, 533,
Uric acid	200-420	100-800	592, 637, 1260, 1368		725, 811, 886, 1136, 1206
Creatinine	25-100			575, 1420, 1459	
Hypoxanthine	8-14	1.8-24.1	740, 1450, 1686	958	725, 1093
Xanthine	0.5-2	0-30	1368, 1686	1250	
Carotenoids			1150, 1520		

Pentru a evidenția contribuția metaboliților purinici în spectrul SERS al biofluidelor, am achiziționat spectrul SERS al unui amestec de acid uric, xantină și hipoxantină – Fig. 4.2. Se poate observa că spectrul SERS al amestecului de metaboliți purinici prezintă caracteristici spectrale similare cu spectrele medii ale fracțiunii metabolice din ser filtrat și urinei, confirmând faptul că metaboliții purinici (în principal acid uric și hipoxantina) aduc principalele contribuții în spectrele SERS ale serului filtrat și urinei. Nu au putut fi identificate benzi SERS specific doar xantinei în spectrul SERS al amestecului celor trei metaboliți purinici, ceea ce indică faptul că afinitatea de suprafață a xantinei este mai mică decât a celorlalți doi metaboliți și majoritatea locurilor de adsorbție disponibile de pe suprafața nanoparticulelor de argint sunt ocupate de acid uric și hipoxantină.

Metaboliții purinici furnizează informații cruciale cu privire la rata de refacere celulară, proces perturbat în aproape fiecare boală. Prin urmare, detectarea anomaliilor maligne se poate realiza prin urmărirea modificărilor semnalului SERS a metaboliților purinici și, implicit, a concentrației metaboliților purinici (xantină, hipoxantină, acid uric) care sunt secretați în cantități ridicate în contextul malignității.

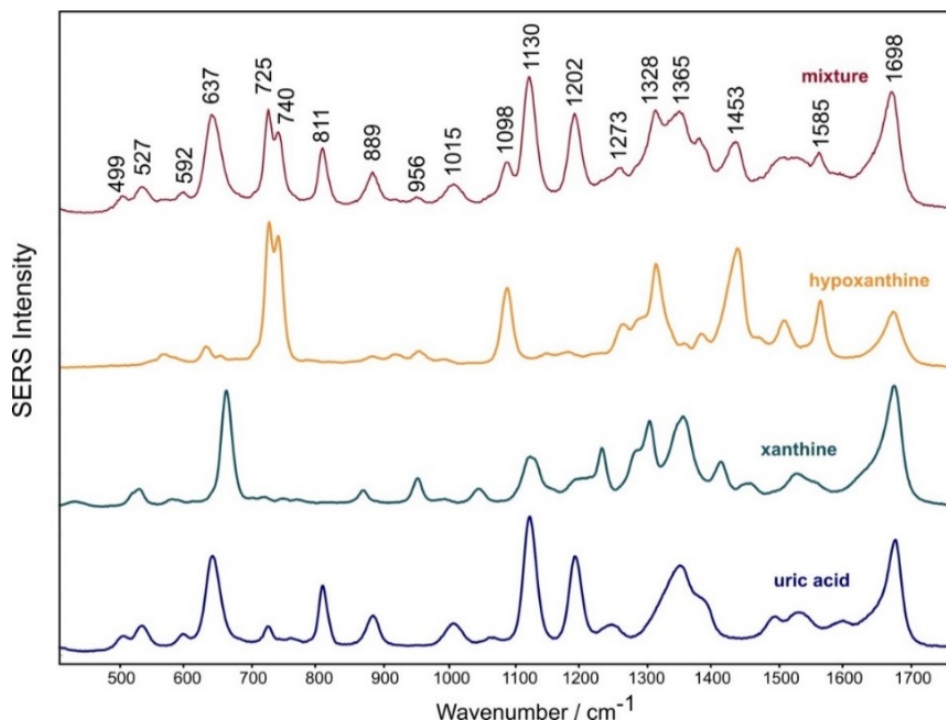


Fig. 4.2 – Spectrele SERS ale acidului uric, xantinei, hipoxantinei și amestecului dintre ele la o concentrație de 0.1 mM fiecare

Carotenoizii sunt o altă clasă de molecule care au putut fi observate numai în spectrele SERS ale fracțiunii metabolice din ser filtrat, având benzi specifice la 1150 cm^{-1} și 1520 cm^{-1} . Spectrele SERS au fost înregistrate cu linia laser de excitație de 532 nm , iar această lungime de undă este în rezonanță cu majoritatea carotenoidelor, prin urmare s-a creat un efect rezonant de împrăștiere a semnalului Raman amplificat la suprafață (SERRS), rezultând un semnal Raman al carotenoidelor mult îmbunătățit. În urină în schimb, s-a observat o contribuție mică a acestora în spectrul SERS (Fig. 4.1). Ca și în cazul metaboliților purinici, carotenoizii au fost sugerați ca posibili markeri de cancer datorită caracterului lor antioxidant, nivelul lor în ser reflectând capacitatea antioxidantă a organismului.

Mai mult decât atât, spectrele SERS ale fracțiunii metabolice din ser filtrat conțin și o bandă la 482 cm^{-1} , caracteristică ergotininei [25], un aminoacid alimentar, care a putut fi detectat doar în probele de ser filtrat din cauza excreției scăzute în urină [45].

Benzile SERS ale creatininei pot fi observate în urină la 575 cm^{-1} și 1420 cm^{-1} [46]. Cunoașterea concentrației de creatinine în urină este importantă în cuantificarea concentrației de proteine din probele de urină (de exemplu, pacienții cu disfuncție renală).

Peak-ul de la 1010 cm^{-1} este atribuit metanolului, folosit pentru a deproteiniza probele de ser filtrat prin precipitarea proteinelor din probă. Pentru banda la 1755 cm^{-1} și umărul la 696 cm^{-1} din probele de ser și benzile la 705 cm^{-1} , 1043 cm^{-1} și 1619 cm^{-1} din probele de urină nu am identificat încă nicio atribuire.

În toate tipurile de cancer, procesul de turn-over celular este accelerat datorită creșterii rapide a metabolismului celular, făcând astfel analiza SERS un bun candidat pentru o metodă point-of-care cu o sensibilitate suficient de mare pentru a detecta malignitatea în stadiu incipient [4, 8, 9, 31, 47, 48]. Cu toate acestea, nivelul ridicat al metaboliților purinici în spectrele SERS ale biofluidelor nu poate fi considerat un biomarker universal specific cancerului. Și alte boli, cum ar fi bolile autoimune [10, 12] prezintă, de asemenea, niveluri crescute de acid uric.

În continuare, pentru a discrimina probele în ser filtrate și urină, am aplicat analiza PCA nesupervizată. Pentru a determina numărul de PC relevante pentru studiul prezent, am analizat varianța pentru fiecare PC în parte și am determinat punctul în care aceasta este maximă, orice creștere peste valoarea respectivă neaducând informații suplimentare importante.

Astfel, am determinat 10 PC-uri, iar în Fig. 4.3 A. se pot observa scorurile componentelor principale și tendința datelor de a se grupa nesupervizat în spațiul creat de PC1 și PC8, care au fost alese automat ca PC-uri optime pentru gruparea PCA nesupravegheată a spectrelor SERS în cele două categorii de biofluide. Scorurile descriu structura datelor în termeni de tipare de eșantion și, în general, arată diferențe sau asemănări ale acestora. Fiecare eșantion are un punctaj pe fiecare PC. Acesta reflectă locația eșantionului de-a lungul PC (coordonata probei de pe PC). Probele cu scoruri apropiate de pe aceeași PC sunt similare (au valori apropiate pentru variabilele corespunzătoare).

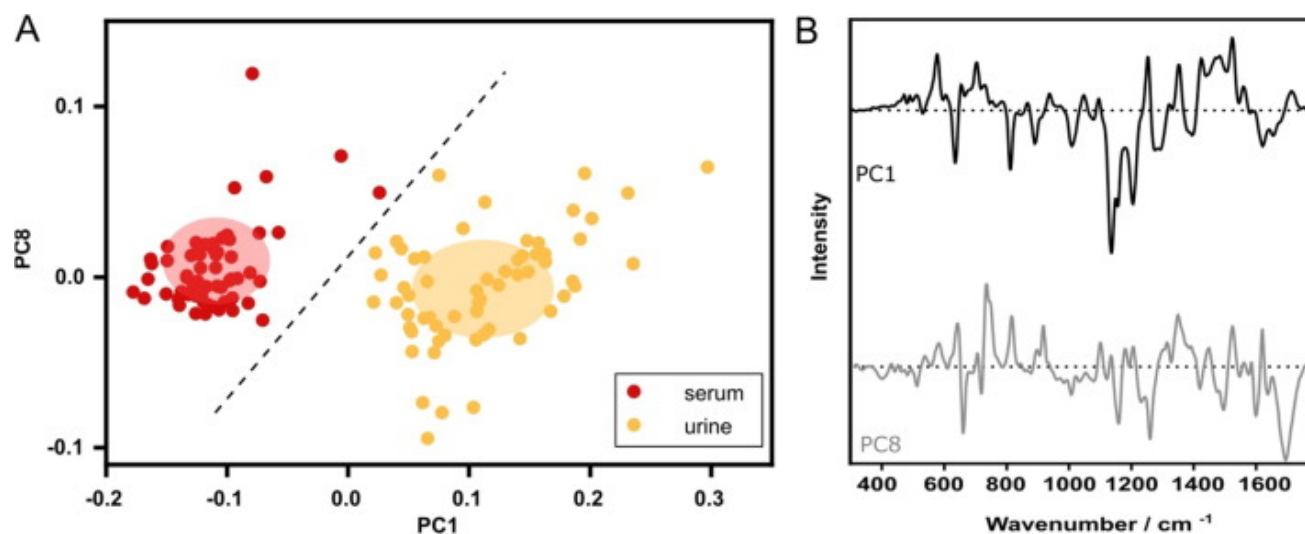


Fig. 4.3 Analiza componentelor principale (PCA) a spectrelor SERS ale serului și urinei

- A.** Score plot-ul PC1 și PC8 – arată gruparea spectrelor SERS ale serului (roșu) și spectrelor SERS ale urinei (galben) **B.** Loading plot-urile corespunzătoare PC1 și PC8

Analizând loading plot-ul corespunzător fiecărui PC (Fig. 4.3 B.), se poate deduce relația dintre scorurile PC, benzile SERS care contribuie cel mai mult la gruparea datelor [49], respectiv similitudinile și diferențele dintre spectrele SERS ale fracțiunii metabolice din ser filtrat și urinei. Se observă că spectrele SERS ale serului sunt grupate în partea negativă a axei PC1, prin urmare, pe baza loading plot-ului corespunzător PC1, putem deduce că probele de ser filtrat au un semnal SERS mai intens de la acidul uric, deoarece benzile SERS ale acidul uric prezintă intensități negative. Benzile de acid uric identificate în loading plot-ul PC1 sunt corelate negativ cu banda SERS a creatininei de la 575 cm⁻¹, deoarece creatinina este detectată numai în probele de urină, ceea ce duce la o grupare nesupravegheată a urinei în partea pozitivă a axei PC1.

Loading plot-ul al PC8 arată benzi SERS pozitive la 725 cm⁻¹ (hipoxantină și acid uric) și 811 cm⁻¹ (acid uric), corelate negativ cu vârfurile de la 648 cm⁻¹, 1150 cm⁻¹, 1250 cm⁻¹ și 1696 cm⁻¹. Se observă o ușoară tendință de grupare a probelor de urină în partea negativă a axei PC8.

4.2. Clasificarea spectrelor SERS în canceros vs necanceros prin algoritmi de machine learning (ML)

Algoritmii de machine learning (ML) sunt utilizate în mod obișnuit pentru clasificarea eșantioanelor bazate pe diferențe subtile în spectrele SERS. Am testat cinci clasificatori ML diferiți (Decision Tree – DT, Random Forest – RF, Gaussian Naïve Bayes – GNB, Support Vector Machine – SVM și Linear Discriminant Analysis – LDA) pentru clasificarea pacienților în două grupuri, cancer de sân – BC și control – CTRL, bazată pe trei seturi variabile: spectrele SERS ale fracțiunii metabolice din ser filtrat, respectiv ale urinei și seturile de date SERS combinate de ser filtrate și urină. Fig. 4.4 ilustrează precizia generală obținută de fiecare dintre cei cinci clasificatori.

LDA este cel mai comun clasificator pentru diagnosticarea cancerului bazată pe biopsia SERS lichidă și într-adevăr, pentru spectrele SERS ale fracțiunii metabolice din ser filtrat, LDA a generat cel mai bun rezultat de clasificare (83%). Cu toate acestea, cea mai mare precizie de clasificare a pacienților în BC și CTRL a fost obținută prin DT (89 %), pe baza spectrelor SERS ale urinei, în timp ce pentru probele de urină LDA a generat o precizie generală de doar 78%, la fel ca RF și GNB.

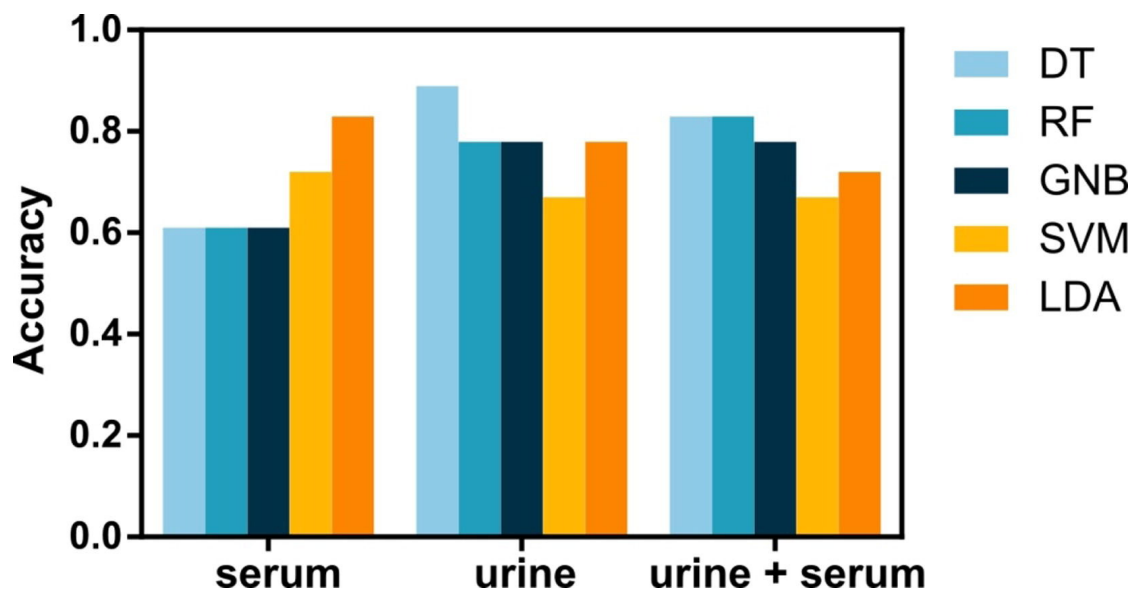


Fig. 4.4 Acuratețea generală a celor cinci algoritmi ML pentru clasificarea probelor în sănătos vs canceros pe baza spectrelor SERS ale serului, urinei și combinării lor

În mod surprinzător, algoritmi ML au oferit o precizie de clasificare mai bună bazată pe spectrele SERS ale urinei în comparație cu spectrele SERS ale fracțiunii metabolice din ser filtrat. Cu toate acestea, prin concatenarea spectrelor SERS de ser filtrat și urină și, ulterior, folosind spectrele SERS combinate pentru clasificarea BC și CTRL, precizia generală nu a fost îmbunătățită, cele mai mari valori fiind atinse de DT și RF (83%). Acuratețea generală, precizia, recall-ul și scorul F_1 pentru evaluarea performanței algoritmilor ML se găsește în Tabelul 4.1.

Tabelul 4.1 Acuratețea generală, precizia, recall-ul și scorul F_1 pentru evaluarea performanței algoritmilor ML

serum	DT	RF	GNB	SVM	LDA
Accuracy	0.61	0.61	0.61	0.72	0.83
Precision	0.73	0.73	0.86	0.70	0.85
Recall	0.67	0.67	0.50	1	0.92
F1-score	0.69	0.69	0.63	0.83	0.88

urine	DT	RF	GNB	SVM	LDA
Accuracy	0.89	0.78	0.78	0.67	0.78
Precision	0.92	0.78	0.83	0.67	0.90
Recall	0.92	0.92	0.83	1	0.75
F1-score	0.92	0.85	0.83	0.8	0.82

combined	DT	RF	GNB	SVM	LDA
Accuracy	0.83	0.83	0.78	0.67	0.72
Precision	0.85	0.85	0.90	0.67	0.82
Recall	0.92	0.92	0.75	1	0.75
F1-score	0.88	0.88	0.82	0.80	0.78

Tabelul 4.2 Acuratețea generală, precizia, recall-ul și scorul F_1 pentru evaluarea performanței analizei PCA combinată cu algoritmi ML

serum	PCA -				
	DT	RF	GNB	SVM	LDA
Accuracy	0.67	0.72	0.83	0.61	0.67
Precision	0.71	0.73	0.80	0.67	0.71
Recall	0.83	0.92	1	0.83	0.83
F1-score	0.77	0.81	0.89	0.74	0.77

urine	PCA -				
	DT	RF	GNB	SVM	LDA
Accuracy	0.61	0.78	0.89	0.83	0.67
Precision	0.73	0.79	0.86	0.85	0.87
Recall	0.67	0.92	1	0.92	0.58
F1-score	0.70	0.85	0.92	0.88	0.70

combined	PCA -				
	DT	RF	GNB	SVM	LDA
Accuracy	0.44	0.72	0.78	0.72	0.67
Precision	0.60	0.73	0.79	0.77	0.87
Recall	0.50	0.92	0.92	0.83	0.58
F1-score	0.55	0.81	0.85	0.80	0.70

În plus, am verificat dacă prin reducerea dimensionalității datelor cu analiza PCA, se îmbunătățește acuratețea clasificării, dar nu s-au găsit îmbunătățiri semnificative ale preciziei de clasificare prin DT, RF, SVM sau LDA după aplicarea PCA (au fost păstrate 20 de PC-uri, ceea ce explică 99% din variația inițială) pentru reducerea dimensionalității – Tabelul 4.2. Singura îmbunătățire semnificativă, se poate observa în cazul GNB, prin creșterea acurateței generale și a scorului F1 atunci când a fost utilizat PCA înainte de clasificare.

Deoarece probele de ser filtrat și urină au fost obținute de la același pacient sau control, am putut verifica rezultatele clasificării obținute cu algoritmul LDA. Fig. 4.5 prezintă spectrele SERS ale probelor fals negative (FN) și fals pozitive (FP), în comparație cu spectrul SERS mediu al grupului din care fac parte. Patru probe au fost clasificate greșit pe baza spectrelor SERS ale fracțiunii metabolice din ser filtrat și trei probe pe baza spectrelor SERS de urină. Cu toate acestea, doar una dintre probele clasificate greșit a fost comună atât pentru probele de ser filtrat, cât și pentru urină (proba 14 din Fig. 4.5 A.), arătând astfel că nu există o corelație între cele două seturi de date și clasificarea greșită.

Am descoperit că probele de ser și urină clasificate greșit au prezentat un profil metabolic anormal în comparație cu grupurile cărora le aparțin, ceea ce a cauzat clasificarea greșită. În special, probele de ser FP și FN au arătat diferențe în intensitatea SERS a benzilor de carotenoide (1150 cm^{-1} și 1525 cm^{-1}) în comparație cu grupul lor original (Fig. 4.5 B.). De asemenea, am observat că probele de ser FN au arătat intensități SERS mai mari ale benzilor de carotenoide (1150 cm^{-1} și 1525 cm^{-1}) în comparație cu media grupului BC, în timp ce nivelurile ridicate de carotenoide au fost mai caracteristice grupului CTRL. În concordanță cu aceste observații, probele de ser filtrat FP au arătat o intensitate mai mică a benzilor de carotenoide, care a fost mai specifică grupului BC. În cazul probelor de urină clasificate greșit, benzile SERS ale creatininei (575 cm^{-1} și 1420 cm^{-1}) au condus la clasificarea greșită a probelor de urină FN și FP (Fig. 4.5 C.).

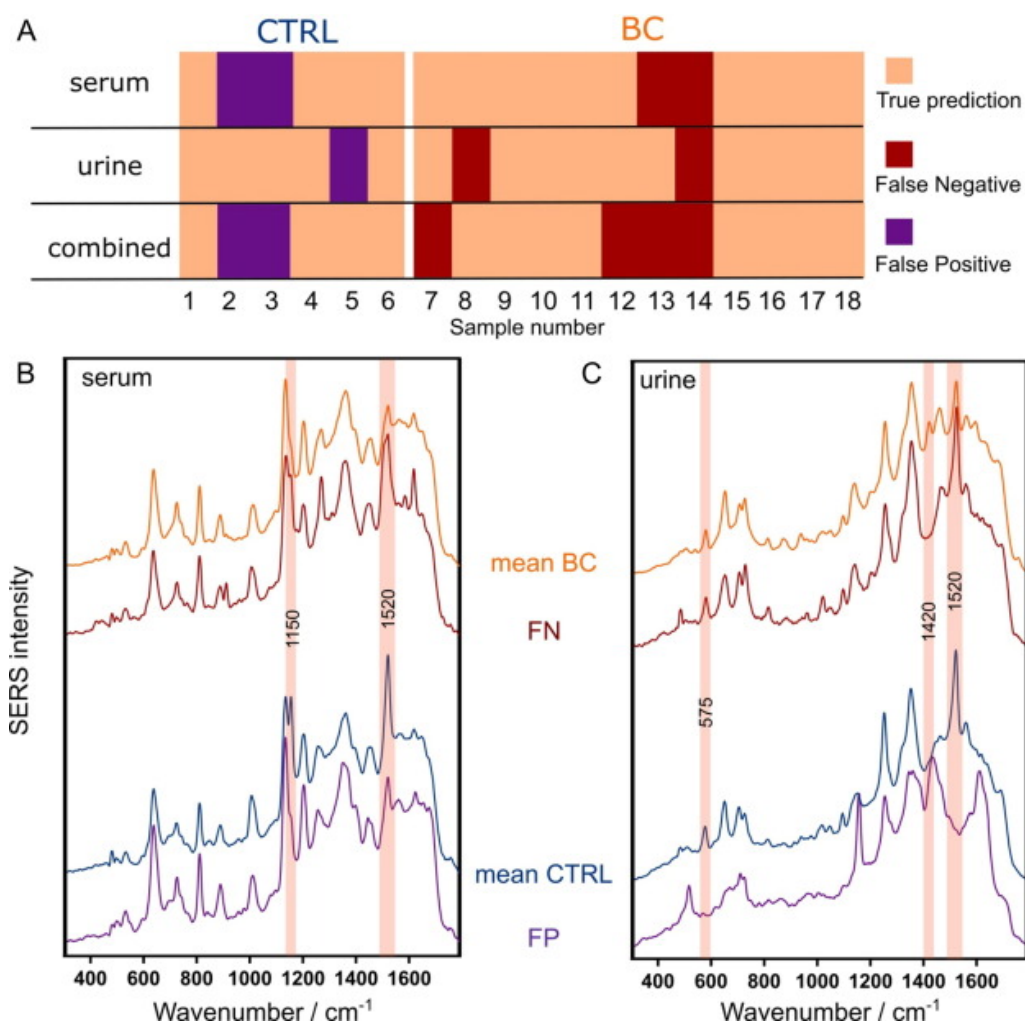


Fig. 4.5 A. Harta cu clasificarea probelor pe baza spectrelor SERS de ser, urină și combinării lor
B. Comparația spectrelor medii SERS ale serului clasificate greșit cu spectrele medii SERS ale grupurilor BC și CTRL
C. Comparația spectrelor medii SERS ale urinei clasificate greșit cu spectrele medii SERS ale grupurilor BC și CTRL

CONCLUZII

Studiul efectuat prezintă aplicabilitatea spectroscopiei SERS în diagnosticul cancerului de sân, împreună cu o metodologie optimă de achiziție. Tehnica de diagnostic pe care se bazează este o biopsie simplă, minim invazivă, analiza realizându-se pe probe de ser filtrat și deproteinizat și urină ale pacientului. Spectrele SERS ale fracțiunii metabolice din ser filtrat și urinei sunt dominate de prezența metaboliților purinici, împreună cu benzi SERS specifice ergotioneinelor și carotenoidelor în probele de ser filtrat și benzi SERS specifice creatininei în probele de urină.

Am testat performanța relativă a cinci algoritmi diferiți de machine learning pentru clasificarea probelor de ser și urină în BC (canceros) vs CTRL (control – necanceros) pe baza spectrelor SERS ale fracțiunii metabolice din ser filtrat și urinei, obținând valori similare în intervalul 61-89%, fapt care indică atât serul filtrat, cât și urina, ca potențiali candidați pentru biopsia lichidă SERS în detectarea cancerului de sân. Precizii de clasificare puțin mai mari au fost obținute utilizând probe de urină în comparație cu probele de ser filtrat, în ciuda variabilității mai mari a spectrelor SERS de urină. Prin identificarea semnăturilor metabolice SERS ale probelor de ser și urină, am putut corela probele clasificate greșit cu dezechilibre specifice în profilul metabolic al probelor respective.

În concluzie, analiza SERS a biofluidelor pentru diagnostic medical este o arie aflată într-o continuă dezvoltare, iar lucrarea de față susține posibilitățile și avantajele aplicabilității spectroscopiei SERS în acest domeniu, putând fi considerată un punct de referință pentru corelarea benzilor SERS cu dezechilibrele metabolice și biochimice în viitoarele biopsii lichide bazate pe SERS.

BIBLIOGRAFIE

1. Ru, E.C.L. and P.G. Etchegoin, *Single-Molecule Surface-Enhanced Raman Spectroscopy*. 2012. **63**(1): p. 65-87.
2. Stefancu, A., S.D. Iancu, and N. Leopold, *Selective Single Molecule SERRS of Cationic and Anionic Dyes by Cl⁻ and Mg²⁺ Adions: An Old New Idea*. The Journal of Physical Chemistry C, 2021. **125**(23): p. 12802-12810.
3. Feng, S., et al., *Nasopharyngeal cancer detection based on blood plasma surface-enhanced Raman spectroscopy and multivariate analysis*. Biosensors and Bioelectronics, 2010. **25**(11): p. 2414-2419.
4. Avram, L., et al., *SERS-Based Liquid Biopsy of Gastrointestinal Tumors Using a Portable Raman Device Operating in a Clinical Environment*. 2020. **9**(1): p. 212.
5. Noothalapati, H., K. Iwasaki, and T. Yamamoto, *Non-invasive diagnosis of colorectal cancer by Raman spectroscopy: Recent developments in liquid biopsy and endoscopy approaches*. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2021. **258**: p. 119818.
6. Vo-Dinh, T., *Biomedical Photonics Handbook*. CRC Press, 2014.
7. Ralbovsky, N.M. and I.K. Lednev, *Towards development of a novel universal medical diagnostic method: Raman spectroscopy and machine learning*. Chemical Society Reviews, 2020. **49**(20): p. 7428-7453.
8. Moisoiu, V., et al., *Breast Cancer Diagnosis by Surface-Enhanced Raman Scattering (SERS) of Urine*. Applied Sciences, 2019. **9**(4): p. 806.
9. Moisoiu, V., et al., *SERS-based differential diagnosis between multiple solid malignancies: breast, colorectal, lung, ovarian and oral cancer*. Int J Nanomedicine, 2019. **14**: p. 6165-6178.
10. Stefancu, A., et al., *SERS-based liquid biopsy of saliva and serum from patients with Sjögren's syndrome*. Anal Bioanal Chem, 2019. **411**(22): p. 5877-5883.
11. Bocsa, C.D., et al., *Knee osteoarthritis grading by resonant Raman and surface-enhanced Raman scattering (SERS) analysis of synovial fluid*. Nanomedicine: Nanotechnology, Biology and Medicine, 2019. **20**: p. 102012.
12. Carlomagno, C., et al., *SERS-based biosensor for Alzheimer disease evaluation through the fast analysis of human serum*. J Biophotonics, 2020. **13**(3): p. e201960033.

13. Chen, S., et al., *Raman Spectroscopy Reveals Abnormal Changes in the Urine Composition of Prostate Cancer: An Application of an Intelligent Diagnostic Model with a Deep Learning Algorithm*. *Advanced Intelligent Systems*, 2021. **3**: p. 2000090.
14. Yang, H., et al., *Noninvasive and prospective diagnosis of coronary heart disease with urine using surface-enhanced Raman spectroscopy*. *Analyst*, 2018. **143**(10): p. 2235-2242.
15. Blanco-Formoso, M. and R.A. Alvarez-Puebla, *Cancer Diagnosis through SERS and Other Related Techniques*. 2020. **21**(6): p. 2253.
16. Derruau, S., et al., *Vibrational Spectroscopy Saliva Profiling as Biometric Tool for Disease Diagnostics: A Systematic Literature Review*. 2020. **25**(18): p. 4142.
17. Chen, S., et al., *Raman Spectroscopy Reveals Abnormal Changes in the Urine Composition of Prostate Cancer: An Application of an Intelligent Diagnostic Model with a Deep Learning Algorithm*. 2021. **3**(4): p. 2000090.
18. Mamede, A.P., et al., *A New Look into Cancer—A Review on the Contribution of Vibrational Spectroscopy on Early Diagnosis and Surgery Guidance*. 2021. **13**(21): p. 5336.
19. Fleischmann, M., P.J. Hendra, and A.J. McQuillan, *Raman spectra of pyridine adsorbed at a silver electrode*. *Chemical Physics Letters*, 1974. **26**(2): p. 163-166.
20. Ferraro, J.R., K. Nakamoto, and C.W. Brown, *Chapter 1 - Basic Theory*, in *Introductory Raman Spectroscopy (Second Edition)*, J.R. Ferraro, K. Nakamoto, and C.W. Brown, Editors. 2003, Academic Press: San Diego. p. 1-94.
21. Stefancu, A., et al., *Combining SERS analysis of serum with PSA levels for improving the detection of prostate cancer*. *Nanomedicine*, 2018. **13**(19): p. 2455-2467.
22. Rickard, J.J.S., et al., *Rapid optofluidic detection of biomarkers for traumatic brain injury via surface-enhanced Raman spectroscopy*. *Nature Biomedical Engineering*, 2020. **4**(6): p. 610-623.
23. Gurian, E., et al., *Differentiation between stages of non-alcoholic fatty liver diseases using surface-enhanced Raman spectroscopy*. *Analytica Chimica Acta*, 2020. **1110**: p. 190-198.
24. Choi, S., et al., *Label-Free Biochemical Analytic Method for the Early Detection of Adenoviral Conjunctivitis Using Human Tear Biofluids*. *Analytical Chemistry*, 2014. **86**(22): p. 11093-11099.
25. Fornasaro, S., et al., *Ergothioneine, a dietary amino acid with a high relevance for the interpretation of label-free surface enhanced Raman scattering (SERS) spectra of many*

- biological samples*. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2021. **246**: p. 119024.
26. Moisoiu, V., et al., *SERS liquid biopsy: An emerging tool for medical diagnosis*. Colloids and Surfaces B: Biointerfaces, 2021. **208**: p. 112064.
 27. Moisoiu, V., et al., *SERS-based differential diagnosis between multiple solid malignancies: breast, colorectal, lung, ovarian and oral cancer*. International journal of nanomedicine, 2019. **14**: p. 6165-6178.
 28. Carmicheal, J., et al., *Label-free characterization of exosome via surface enhanced Raman spectroscopy for the early detection of pancreatic cancer*. Nanomedicine: Nanotechnology, Biology and Medicine, 2019. **16**: p. 88-96.
 29. Chundayil Madathil, G., et al., *A Novel Surface Enhanced Raman Catheter for Rapid Detection, Classification, and Grading of Oral Cancer*. Advanced Healthcare Materials, 2019. **8**(13): p. 1801557.
 30. Davis, R.M., et al., *Surface-Enhanced Raman Scattering Nanoparticles for Multiplexed Imaging of Bladder Cancer Tissue Permeability and Molecular Phenotype*. ACS Nano, 2018. **12**(10): p. 9669-9679.
 31. Hong, Y., et al., *Label-free diagnosis for colorectal cancer through coffee ring-assisted surface-enhanced Raman spectroscopy on blood serum*. Journal of Biophotonics, 2020. **13**(4): p. e201960176.
 32. Moisoiu, V., et al., *Breast Cancer Diagnosis by Surface-Enhanced Raman Scattering (SERS) of Urine*. Vol. 9. 2019. 806.
 33. Iancu, S.D., et al., *The role of Ag(+), Ca(2+), Pb(2+) and Al(3+) adions in the SERS turn-on effect of anionic analytes*. Beilstein J Nanotechnol, 2019. **10**: p. 2338-2345.
 34. Leopold, N., et al., *The role of adatoms in chloride-activated colloidal silver nanoparticles for surface-enhanced Raman scattering enhancement*. Beilstein J Nanotechnol, 2018. **9**: p. 2236-2247.
 35. Stefancu, A., et al., *Specific and selective sers active sites generation on silver nanoparticles by cationic and anionic adatoms*. Romanian Reports in Physics, 2018. **70**(4).
 36. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries*. CA Cancer J Clin, 2018. **68**(6): p. 394-424.

37. Ashikari, R., et al., *Paget's disease of the breast*. *Cancer*, 1970. **26**(3): p. 680-5.
38. Begg, C.B., et al., *Variation of breast cancer risk among BRCA1/2 carriers*. *Jama*, 2008. **299**(2): p. 194-201.
39. Kösters, J.P. and P.C. Göttsche, *Regular self-examination or clinical examination for early detection of breast cancer*. *The Cochrane database of systematic reviews*, 2003(2): p. CD003373.
40. Bonifacio, A., et al., *Surface-enhanced Raman spectroscopy of blood plasma and serum using Ag and Au nanoparticles: a systematic study*. *Anal Bioanal Chem*, 2014. **406**(9-10): p. 2355-65.
41. Leopold, N. and B. Lendl, *A New Method for Fast Preparation of Highly Surface-Enhanced Raman Scattering (SERS) Active Silver Colloids at Room Temperature by Reduction of Silver Nitrate with Hydroxylamine Hydrochloride*. *Journal of Physical Chemistry B - J PHYS CHEM B*, 2003. **107**.
42. Haiss, W., et al., *Determination of Size and Concentration of Gold Nanoparticles from UV-Vis Spectra*. *Analytical Chemistry*, 2007. **79**(11): p. 4215-4221.
43. rafael; Scott M Rosendahl; Rahul Kumar, M.T.S.R.b.J.E.L.J.o.V.B.I.N.A.a.P.G.r.A.L.S., *Quasars/orange-spectroscopy: Release 0.5.8 (Version 0.5.8)*. *Zenodo*. 2021.
44. Westley, C., et al., *Absolute Quantification of Uric Acid in Human Urine Using Surface Enhanced Raman Scattering with the Standard Addition Method*. *Analytical Chemistry*, 2017. **89**(4): p. 2472-2477.
45. Tang, R.M.Y., et al., *Distribution and accumulation of dietary ergothioneine and its metabolites in mouse tissues*. *Scientific Reports*, 2018. **8**(1): p. 1601.
46. Li, M., et al., *Reagent- and separation-free measurements of urine creatinine concentration using stamping surface enhanced Raman scattering (S-SERS)*. *Biomedical Optics Express*, 2015. **6**(3): p. 849-858.
47. Cervo, S., et al., *SERS analysis of serum for detection of early and locally advanced breast cancer*. *Anal Bioanal Chem*, 2015. **407**(24): p. 7503-9.
48. Del Mistro, G., et al., *Surface-enhanced Raman spectroscopy of urine for prostate cancer detection: a preliminary study*. *Anal Bioanal Chem*, 2015. **407**(12): p. 3271-5.
49. *Principal Component Analysis for Special Types of Data*, in *Principal Component Analysis*, I.T. Jolliffe, Editor. 2002, Springer New York: New York, NY. p. 338-372.

