

# **Prediktív biomarkerek vizsgálata különböző daganattípusokban**

Doktori tézisek

**Fekete János Tibor**

Semmelweis Egyetem

Patológiai tudományok Doktori Iskola



Témavezető:

Dr. Györfly Balázs D.Sc., egyetemi tanár

Hivatalos bírálók:

Dr. Vas Virág Ph.D., tudományos főmunkatárs

Dr. Borka Katalin Ph.D., egyetemi docens

Szigorlati bizottság elnöke:

Dr. Szijártó Attila D.Sc., egyetemi tanár

Szigorlati bizottság tagjai:

Dr. Monostory Katalin Ph.D., tudományos főmunkatárs

Dr. Fekete Andrea Ph.D., egyetemi docens

Budapest

2021

## 1. Bevezetés

Dolgozatom elsődleges témája a prediktív biomarkerek vizsgálata volt különböző daganattípusokban. A prediktív biomarkerek a terápiás beavatkozások kimenetelének előrejelzésére alkalmasak, ennek megfelelően szolgáltathatnak információt a terápia várható sikerességéről, a terápiás rezisztencia meglétéről, vagy a súlyos mellékhatás kialakulásáról.

Az emlőtumorok esetében például az ösztrogén, a progeszteron, illetve a HER2 receptorok kifejeződése nem csak prognosztikus, de prediktív, a gyógyszeres terápiát meghatározó információt is hordoznak. Ugyan az elmúlt években több, a klinikumban használt multigénese eszközt is engedélyeztek, ezek jellemzően prognosztikus információkat nyújtanak. Az emlőtumorok génexpresszió alapú prediktív biomarkereinek a vizsgálatait segítő online alkalmazás a további biomarker kutatásokhoz kíván segítséget nyújtani.

A petefészekdaganatos betegek gyógyszeres kezelésének arany standardja a platina és taxán kombinált kezelés. Habár a terápiás választ adók aránya a terápia kezdetén elérheti a betegek 70-80%-át is, a tumor kiújulási rátája magas. A kemoterápia kimenetelét több tényező, így a kezelés elleni rezisztencia mechanizmusok kialakulása is befolyásolja, amelyre több gén expressziós módosulása is hatással van. Dolgozatomban a platina és taxán kezeléssel szembeni kemorezisztenciában potenciális szerepet betöltő gének azonosításával és azok klinikai mintákon történő igazolásával foglalkoztam.

A mikro-RNS-ek (miRNS) olyan rövid, 22-25 nukleotid bázisból álló, nem kódoló RNS molekulák, amelyek a gének transzkripciót követő szabályozásában vesznek részt. A miRNS szint módosulásai különböző betegségek patogenezisében is szerepet játszanak, ezen felül több kísérletes vizsgálat is bizonyította, hogy a terápiás válasz sikerességét is befolyásolják. Munkámban a platina alapú kemoterápia kimenetelét befolyásoló miRNS mintázatok azonosításával foglalkoztam három tumortípusban.

A daganatos megbetegedések előfordulása és a betegség okozta halálozások száma világszerte növekvő tendenciát mutat. Az Egészségügyi Világszervezet adatai alapján a daganatok a 70 éven aluliak csoportjában az első vagy második vezető halálokot jelentik a statisztikában szereplő 183 országból 112-ben. Hazánk mutatói különösen kedvezőtlenek. Munkánkban az elmúlt évek incidenciája és mortalitási összefoglalóján túl arra kerestük a választ, hogy a közeli jövőben milyen tendenciák várhatók hazánkban.

## **2. Célkitűzés**

A munkám során a célkitűzéseim a következők voltak az egyes témákkal kapcsolatban:

2.1. Prediktív biomarkerek azonosítására és validálására alkalmas online alkalmazás készítése emlődaganatos minták génexpressziós adatainak felhasználásával.

2.2. A petefészekdaganatok génexpresszió alapú potenciális prediktív biomarkereinek azonosítása és ezek igazolása független klinikai betegmintákon is.

2.3. A platina alapú kezelésben részesült laphámkarcinómás esetek miRNS mintázatának azonosítása és gépi tanulással támogatott elemzése.

2.4. A rosszindulatú daganatos megbetegedések hazánkban mért incidenciájának és mortalitásának dokumentálása az 1996-2015 közti időszakban, illetve előrejelzés készítése a 2025 és 2030-as évekre.

### **3. Módszerek**

#### **3.1. mRNS és miRNS expressziós adatok feldolgozása**

Az mRNS expressziós adatok és az esetekhez tartozó klinikai annotáció forrásaként több adattárházat – Gene Expression Omnibus (GEO), Array Express, és TCGA / GDC (The Cancer Genome Atlas / Genomic Data Commons) is használtam.

A forráskeresést követően mindegyik munkám esetében egy aggregált adatbázist hoztam létre, amelynek kialakítása magában foglalta az adatok minőségellenőrzését, a normalizálását, valamint az adattisztítást is. Az így létrejött adatbázisokat használtam az elemzések forrásaként.

A szenzitív és rezisztens betegcsoportok kialakítására alapvetően két módszert alkalmaztam. Az emlő és a petefészek tumoros mintákkal foglalkozó munkáimban ehhez a patológias

komplett választ, illetve a túlélési információkat (relapszus, illetve progresszió mentes túlélés) használtam fel, míg a laphámkarcinómás vizsgálatban a szenzitív és rezisztens csoportosítás a medián progressziómentes túlélésen alapult.

### **3.2. Epidemiológia adatok forrása és feldolgozása**

A vizsgálathoz a Központi Statisztikai Hivatal biztosította a mortalitási adatokat az 1996-2015 közötti időszakra, valamint a demográfiai előrejelzést. Az incidencia adatok vizsgálatára a Nemzeti Rákregiszter honlapjáról letöltött 2001-2014 közötti adatsorokat használtuk fel.

Az elemzés során standardizálást leíró statisztikát és regressziós modelleket használtunk.

### **3.3. Statisztikai számítások**

A prediktív biomarkerekkel kapcsolatos munkáim során az egyes gének és miRNS-ek kifejeződésének az összehasonlításához a szenzitív és rezisztens betegcsoportok között Mann-Whitney U tesztet, valamint Receiver Operating Characteristic (ROC) elemzést használtam.

A laphámkarcinómák miRNS mintázatának vizsgálata során a Mann-Whitney U teszten és a ROC elemzésen túl, fürtelemzést és korrelációanalízist végeztem az egyes tumortípusok közötti hasonlóság meghatározásához. A platina terápia kimenetelének előrejelzésére szolgáló modell a méhnyak és fej-nyaki tumorerő aggregált adatbázisában logisztikus regresszióval készült. Az egyes miRNS-ek potenciális célgénjeinek meghatározása a DIANA-mirPath 3.0 webalkalmazás felhasználásával történt.

A statisztikai számítások R statisztikai programkörnyezetben Bioconductor programcsomagok felhasználásával történt. A statisztikai szignifikancia határértékeként  $p < 0,05$  került beállításra.

Azokban az esetekben, ahol többszörös hipotézistesztelet végeztem p-érték korrekciót is végeztem a hamis találati arány (FDR) megadásával.

### **3.4. Klinikai minták**

Az *in vitro* validációra használt petefészek daganatos betegmintákat az Országos Onkológiai Intézet biztosította. A minták génexpressziós mintázatának vizsgálata kvantitatív polimeráz reakcióval történt.

## **4. Eredmények**

### **4.1. Prediktív biomarkerek azonosítása emlőtumorokban**

Az online alkalmazás adatbázisának építéséhez az NCBI GEO és Array Express adattárházak adatait használtam fel, amely összességében így 3.104 betegminta génexpressziós adatait tartalmazza.

Az adatbázis validálása két lépésben történt. Első körben az elemzést olyan ismert biomarkereken végeztem el, amelyek expressziója jellemző a betegség bizonyos altípusaiban. A HER2 gén trastuzumab kezelés esetén ( $p=8,4E-04$ ,  $AUC=0,629$ ), míg a PGR gén a hormonterápia esetén ( $p=8,6E-05$ ,  $AUC=0,7$ ) bizonyult a terápia sikerességét előrejelezni képes biomarkernek. Második validációs lépésként a paclitaxel rezisztenciában potenciálisan szerepet játszó géneket

ellenőriztem. A rezisztenciával kapcsolatba hozott 29 gén mindegyikét ellenőriztem és összesen 23 esetben találtam szignifikáns összefüggést a terápiás válasszal.

Végül, valamennyi olyan tripla-negatív altípusba sorolt, a kezelés során kemoterápiában részesült betegmintán lefuttattam az elemzést, akik esetében a patológiás komplett választ használtuk végpontként (n=473). Az eredményeket ROC AUC szerint rendezve a legjobb három gén a FUS ( $p = 3,30E-16$ , AUC = 0,699), a LMNB1 ( $p = 2,20E-16$ , AUC = 0,698) és az SNRNP40 ( $p = 4,20E-15$ , AUC = 0,694) voltak.

Az online alkalmazás nyílt forráskódú PHP programnyelven készült, az adatokat PostgreSQL adatbázisban tároljuk. A webfelület és a háttéradatbázis közötti kapcsolat R szerver biztosítja. Az alkalmazás a [www.rocplot.org](http://www.rocplot.org) címen érhető el.

## **4.2. A platina-taxán kezelés biomarkerei petefészek daganatokban**

A vizsgálat első részében adatbázist építettem, amely 1.816 petefészekdaganatos betegminta génexpressziós adatait, valamint klinikai annotációját tartalmazza. Az adatbázist használva bioinformatikai elemzést végeztem olyan serosus szövettani típusú petefészekdaganatos mintákon (n=871), akik a kezelés során platina és taxán terápiában részesültek. Az elemzést követően a klinikai mintákon történő validációra nyolc gént választottam ki - AKIP1, AKIRIN2, MARVELD1, CFL1, SERBP1, PDXK, TFE3 és NCOR2 -, amelyek mind a ROC elemzés, mind pedig a Mann-Whitney elemzés során a legszignifikánsabb összefüggést mutatták a terápiás válasszal.

A fenti gének esetében kvantitatív PCR elemzést végeztem az Országos Onkológiai Intézet által biztosított betegmintákon, amely hat esetben – AKIP1, AKIRIN2, MARVELD1, PDXK, TFE3 és NCOR2 - igazolta a bioinformatikai elemzés eredményét. A kiválasztott gének mindegyikének a kifejeződése a rezisztens betegmintákban volt magasabb.

### **4.3. A terápiás választ meghatározó miRNS mintázatok laphámkarcinómákban**

A laphámkarcinómák terápiás választ meghatározó miRNS mintázatának meghatározásához a GDC adattárházból töltöttem le és dolgoztam fel olyan fej-nyaki (105 beteg), méhnyak (94 beteg) és tüdő (67 beteg) tumoros esetek klinikai és miRNS expressziós adatait, akik terápiájuk során platina kezelésben részesültek. A minták kategorizálása szenzitív és rezisztens csoportokra a progressziómentes túlélésen alapult (18 hónap).

Az egyes tumortípusokat először külön vizsgálva méhnyak tumorok esetében 16, fej-nyaki tumorok esetében 103, míg tüdő laphámtumorok esetében kilenc olyan miRNS-t azonosítottam, amelyek kifejeződése meghatározó volt a terápiás válasz kimenetelében. A 128 szignifikáns miRNS közül 11 két tumortípusban is megegyezett, amelyből három a fej-nyaki és tüdő, nyolc pedig a fej-nyaki és a méhnyak tumorok esetében volt közös. Az egyes tumortípusok miRNS kifejeződésének hasonlósági vizsgálata azt mutatta, hogy a terápiás választ meghatározó mintázat a fej-nyaki és méhnyak daganatos minták esetében mutat hasonlóságot. A két tumortípust együtt vizsgálva gépi tanulással, a logisztikus regressziós modell hat olyan



miRNS-t azonosított (hsa-miR-5586, hsa-miR-632, hsa-miR-2355, hsa-miR-642a, hsa-miR-101-2, hsa-miR-6728), amelyek kifejeződése szignifikáns kapcsolatot mutat a platina terápia sikerességével.

A modell által meghatározott miRNS-ek célgénjei a purin és pirimidin metabolizmusban, illetve a cirkadián szabályozásban vesznek részt.

#### **4.4. A daganatos betegségek incidenciája és mortalitása Magyarországon az 1996-2030 időszakban**

Munkánk célja az volt, hogy egyrészt áttekintsük az 1996-2015 közötti periódusban regisztrált újonnan diagnosztizált, valamint a betegséghez köthető halálozási esetszámokat tumortípusonként, nemenként és korcsoportonként. Másrészt, hogy ezen adatokra támaszkodva olyan statisztikai modellt készítsünk, amely a várható demográfiai változásokat is figyelembe véve becsüli meg a hazánkban várható rákteher mértékét 2025-re és 2030-ra.

A 2006-2015-ös időszakban 712 ezer új esetet diagnosztizáltak hazánkban. Férfiak esetében a leggyakrabban azonosított daganatok a tüdő-, a vastagbél és végbél-, valamint a prosztaták voltak. Ez a három típus az összes regisztrált eset 45%-a. Nőknél az emlő, a kolorektális és a tüdő daganatos megbetegedései az összes eset 46 százalékáért voltak felelősek.

Ugyanezt az időszakot vizsgálva 331 ezer daganatos megbetegedéshez köthető halálozást regisztráltak. A férfiak körében a tüdő és a kolorektális daganatok okozta halálozás az összes haláleset 45%-át tette ki. A nők körében a tüdődaganatok

(20%), a kolorektális daganatok (14%) és az emlődaganatok (14%) jelentették a leggyakoribb halálokat.

A rendelkezésünkre álló adatok alapján 2016 és 2030 között hazánk lakossága várhatóan 6%-kal fog zsugorodni a 2015-ös állapothoz viszonyítva. A daganatos megbetegedések egyik jelentős rizikófaktora az életkor. Ebből következően, a fiatalabb korosztályokban várható létszámcsökkenés egyben azt is jelenti, hogy a daganatfejlődés szempontjából kisebb relatív kockázatot viselő korcsoportok aránya is csökken a teljes népességben belül.

Az incidenciát vizsgálva megállapítható, hogy a vizsgált időszakban (1996-2030) az európai populáció korösszetételét használva a standardizáláshoz, a 100.000 főre vetített arányszám férfiakban enyhén csökkenő, míg a nők esetében folytonosan növekvő trendet mutat, míg a nyers arányszámmal vizsgálva mindkét nem esetében növekedés figyelhető meg. A mortalitás esetében az európai populációhoz standardizált 100.000 főre vetített arányszám csökkenő, míg a nyers ráta növekvő trendet mutat mindkét nemben.

A nyers incidenciá arányszámok tekintetében a 2001-2030 időszakban a férfiak esetében az átlagos éves százalékos változás (annual percentage change, APC) 1,35%, míg a nők esetében ez 1,68%. A nyers halálozási arányszámok esetében az 1996-2030 közötti időszakban a férfiaknál 0,78%, míg nőknél 1,47%-os az átlagos éves százalékos változás mértéke.

Férfiak esetében a vastagbél és végbél daganatoknál kalkulált évi 2,1%-os növekedés következtében várható, hogy az esetszám meghaladja a tüdődaganattal diagnosztizáltak számát az előre jelzett periódus végére (2030). Habár a tüdődaganatok

okozta halálozás várhatóan vissza fog szorulni, még így is a leggyakoribb tumor okozta elhalálozásért felelős típus marad a 40-79 éves korosztályban.

A nők között becslésünk szerint mind az incidenciát, mind pedig a mortalitást vizsgálva az emlő-, a vastagbél és végbél, valamint a tüdő daganatos megbetegedései lesznek a meghatározó onkológiai betegségek 2030-ban is.

## **5. Következtetések**

5.1. A leggyakoribb kemoterápiás, valamint a célzott terápiás kezelések génexpresszióan alapuló prediktív biomarkereinek azonosítására és validálására alkalmas online alkalmazás lehetőséget biztosít az egyes gének szerepének vizsgálatára emlőtumoros mintákon. Az alkalmazás validálása során egy korábbi vizsgálat 29 terápia rezisztenciával összefüggésbe hozott gén közül 23 gént sikerült igazolni.

5.2. A petefészek daganatokban bioinformatikai elemzéssel azonosított szignifikáns gének független, klinikai mintákon történő validálása során hat, a platina és taxán kombinált terápia sikerességét előrejelezni képes gént igazoltam. A hat gén (AKIP1, AKIRIN2, MARVELD1, NCOR2, PDXK, TFE3) magasabb expressziója szerepet tölthet be a kemoterápia rezisztencia kialakulásában.

5.3. A laphámkarcinómás esetek feldolgozása során azonosítottam a méhnyak, fej-nyaki és tüdő tumorokban,

valamint az egyes tumortípusok között átfedést mutató miRNS-eket, amelyek befolyásolhatják a platina alapú kezelés kimenetelét. A gépi tanulás segítségével azonosított hat miRNS-t tartalmazó mintázat hatékonyan volt képes előrejelezni a platina alapú terápia sikerességét méhnyak és fej-nyaki tumorok esetében.

5.4. A Központi Statisztikai Hivatal és a Nemzeti Rákregiszter 1996-2015 közötti adatainak felhasználásával leírásra került a rosszindulatú daganatok incidenciája és mortalitása. A korábbi időszak eredményeit, valamint a demográfiai mutatókat figyelembe vevő modell a várható rákterhelés előrejelzésével segíthet a prevenció programok, szűrések és a kutatási irányok tervezésében.

## 6. Saját publikációk jegyzéke

### 6.1. Disszertációhoz kapcsolódó publikációk

- **Fekete JT**, Győrffy B. ROCplot.org: Validating predictive biomarkers of chemotherapy/hormonal therapy/anti-HER2 therapy using transcriptomic data of 3,104 breast cancer patients. *Int J Cancer*. 2019 Dec 1;145 (11): 3140-3151. **IF: 5,145**
- **Fekete JT**, Ősz Á, Pete I, Nagy GR, Vereczkey I, Győrffy B. Predictive biomarkers of platinum and taxane resistance using the transcriptomic data of 1816 ovarian cancer patients. *Gynecol Oncol*. 2020 Mar;156(3):654-661. **IF: 4,623**
- **Fekete JT**, Welker Á, Győrffy B. miRNA Expression Signatures of Therapy Response in Squamous Cell Carcinomas. *Cancers (Basel)*. 2020 Dec 28;13(1):63. **IF: 6,126**
- Menyhárt O, **Fekete JT**, Győrffy B. Demographic shift disproportionately increases cancer burden in an aging nation: current and expected incidence and mortality in Hungary up to 2030. *Clin Epidemiol*. 2018 Aug 29;10:1093-1108. **IF: 3,178**

## 6.2. Disszertációtól független publikációk jegyzéke

- Menyhárt O, **Fekete JT**, Győrffy B. Gene Expression Indicates Altered Immune Modulation and Signaling Pathway Activation in Ovarian Cancer Patients Resistant to Topotecan. *Int J Mol Sci.* 2019 Jun 5;20(11):2750. **IF: 4,556**
- Pérez-Pena J, **Fekete JT**, Páez R, Baliu-Piqué M, García-Saenz JÁ, García-Barberán V, Manzano A, Pérez-Segura P, Esparis-Ogando A, Pandiella A, Győrffy B, Ocana A. A Transcriptomic Immunologic Signature Predicts Favorable Outcome in Neoadjuvant Chemotherapy Treated Triple Negative Breast Tumors. *Front Immunol.* 2019 Dec 18;10:2802. **IF: 5,085**
- Menyhárt O, **Fekete JT**, Győrffy B. Gene expression-based biomarkers designating glioblastomas resistant to multiple treatment strategies. *Carcinogenesis.* 2021 Mar 23:bgab024. **IF: 4,603**
- Lehotzky A, Oláh J, **Fekete JT**, Szénási T, Szabó E, Győrffy B, Várady G, Ovádi J. Co-Transmission of Alpha-Synuclein and TPPP/p25 Inhibits Their Proteolytic Degradation in Human Cell Models. *Front Mol Biosci.* 2021 May 18; 8:666026. **IF: 4,188**