

**Egyetemi doktori (PhD) értekezés tézisei**

**KÖNNYPROTEOMIKAI ÉS KÉPFELDOLGOZÁSI TECHNIKÁK  
ALKALMAZHATÓSÁGÁNAK VIZSGÁLATA A RETINOPATHIA  
DIABETICA SZŰRÉSÉBEN**

Dr. Török Zsolt

Témavezető: Dr. Csutak Adrienne



DEBRECENI EGYETEM

Klinikai Orvostudományok Doktori Iskola

Debrecen, 2015.

# **KÖNNYPROTEOMIKAI ÉS KÉPFELDOLGOZÁSI TECHNIKÁK ALKALMAZHATÓSÁGÁNAK VIZSGÁLATA A RETINOPATHIA DIABETICA SZŰRÉSÉBEN**

Értekezés a doktori (PhD) fokozat megszerzése érdekében klinikai  
orvostudományok tudományágban.

Írta: Dr. Török Zsolt okleveles orvos

Készült a Debreceni Egyetem Klinikai orvostudományok doktori iskolája  
(Klinikai vizsgálatok programja) keretében.

Témavezető: Dr. Csutak Adrienne, PhD

A doktori szigorlati bizottság:

elnök: Prof. Dr. Hernádi Zoltán, az MTA doktora  
tagok: Prof. Dr. Németh János, az MTA doktora  
Prof. Dr. Paragh György, az MTA doktora

A doktori szigorlat időpontja: Debreceni Egyetem ÁOK, Szülészeti és  
Nőgyógyászati Intézet tanterme, 2015. március 12. 11 óra.

Az értekezés bírálói: Dr. Bereczky Zsuzsanna, PhD  
Dr. Hidasi Vanda PhD

A bírálóbizottság:

elnök: Prof. Dr. Hernádi Zoltán, az MTA doktora  
tagok: Prof. Dr. Németh János, az MTA doktora  
Prof. Dr. Paragh György, az MTA doktora  
Dr. Bereczky Zsuzsanna, PhD  
Dr. Hidasi Vanda PhD

Az értekezés védésének időpontja: Debreceni Egyetem ÁOK, Belgyógyászati  
Intézet, „A” épület tanterme, 2015. 03. 12. 13 óra.

# 1. Előzmények

## Bevezetés

Munkánk a diabetes mellitus (DM) szövődményeként fellépő ***retinopathia diabetica (RD) szűrésére alkalmazható automatizálható módszertanok vizsgálatára irányult.***

A téma fontosságát az alapbetegség, illetve szövődményeinek magas prevalenciája adja, valamint az a tény, hogy a RD-t, az aktív életkorban bekövetkező látásvesztés leggyakoribb kiváltó okaként tartják nyilván.

A RD kezdetben tünetmentes, így a betegek sokszor csak igen későn, a betegség előrehaladott stádiumában fordulnak orvoshoz. A retinopathia progressziója korai diagnózissal és kezeléssel jelentősen lassítható, ezért fontos a veszélyeztetett betegcsoport – a DM-ban szenvedő betegek – rendszeres szemészeti szűrővizsgálata.

A RD szűrést napjainkban egészségügyi képesítéssel bíró szakemberek végzik, akik a legtöbb esetben személyesen találkoznak a betegekkel a vizsgálat során. A központosított szűrővizsgálatok esetén a vizsgáló csak a beteg fundusáról készített digitális fotót kapja meg és használja az

értékeléshez. Mindkét protokoll hátránya, hogy alkalmazásához jelentős számú, speciálisan képzett szakemberre van szükség, aminek költségvonzata igen magas.

Az utóbbi években számos kezdeményezés indult útjára azzal a céllal, hogy a RD szűrése során hagyományosan alkalmazott protokollokba olyan módszertani lépéseket illesszenek, melyek – támaszkodva a modern informatikai és laboratóriumi technológiákra – csökkentik a szűrővizsgálatok humánerőforrás igényét. A humánerőforrás igény csökkentése a szűrőprogramok költségének csökkenésével járhat, ami a vizsgálatokat a DM-ban szenvedő betegek szélesebb köre számára teheti elérhetővé.

Munkánk során ezen új módszertanok közül az alábbiakat vizsgáltuk:

- ***a könnyben található fehérjék RD hatására bekövetkező, jellemző koncentráció változásának vizsgálata proteomikai eljárással***
- ***a digitalizált szemfenéki képeken azonosítható, RD-ra jellemző elváltozások automatizált felismerése a képfeldolgozás eszközeivel***
- ***a két fenti módszertan kombinálására épülő eljárások.***

## **A retinopathia diabetica kialakulása**

A DM globális prevalenciáját a 20–79 éves korcsoportban 6,4%-osra becsülték 2010-ben, ami 285 millió beteget jelentett. Egy 1980 és 2008 között készült, 35 tanulmány adatait feldolgozó munka alapján a vizsgált 22.896 diabeteses betegre vonatkozóan a RD teljes prevalenciája 34,6%-volt. A 40-60 éves korosztályban a leggyakoribb vaksághoz vezető okok a rövidlátás (myopia) és a RD, 60 év felett a maculadegeneratio és a RD voltak.

A RD progressziója során, a különböző szemfenéki elváltozások megjelenése alapján két stádiumot különböztetnek meg: i.) Nonproliferatív (enyhe, mérsékelt és súlyos altípus); ii.) Proliferatív retinopathia (enyhe-közepes és magas rizikójú alcsoport). Maculopathia mind a nonproliferatív, mind a proliferatív stádiumban megjelenhet. A szemfenéki elváltozások között a képfeldolgozási eszközökkel történő detektálás szempontjából legfontosabb a nonproliferatív stádiumban megjelenő, szemtükörrel apró vörös pontként látható microaneurysma (MA). Az 1. stádiumban emellett intraretinális bevérzések, kemény exsudatumok és a macula oedema is megjelenhet (2. ábra). A pont/folt alakú bevérzések a capillarisokrupturája következtében alakulnak ki és többnyire a MA-nál mélyebb rétegben helyezkednek el. A kemény exsudatumokat a megnövekedett membrán permeabilitás

következtében kiszivárgó lipidek lerakódása hozza létre. Jelen munka szempontjából kiemelt szerepe van a MA-nak, mivel ez a laesio a RD legkorábbi jelei közé tartozik. Ha a fundus vizsgálata során MA-t nem tudunk azonosítani a RD diagnózisa megkérdőjelezhető.

## **A jelenleg alkalmazott RD szűrési módszerek**

Az egyik legeredményesebb szűrőprogram az „*English National Screening Programme for Diabetic Retinopathy*”, ahol – standardok szerint képzett – szakemberek különítik el az egészséges és a RD tüneteit mutató digitalizált retinaképeket, majd küldik további vizsgálatokra azon pácienseket, akiknél súlyosabb fokú retina elváltozást találtak. A digitális képek értékelését végző szakemberek gyakran a betegvizsgálat helyszínétől távol, klinikai leolvasó centrumokba tömörülnek, ahonnan aleleteket néhány nap elteltével küldik vissza a vizsgálat kérőinek. A RD szűrővizsgálatoknak alapvetően kétféle módszertani megközelítése létezik:

- fundus vizsgálat: direkt vagy indirekt szemtükrözéssel,
- retinafényképezés a hozzá kapcsolódó osztályozással.

A *Health Technology Board for Scotland* megállapítása szerint a réslámpák használatával végzett indirekt szemtükrözés érzékenysége és specificitása alapján megfelelő eszköz a retinaszűrés elvégzésére, azonban, mivel a folyamat során nem készül fotódokumentáció a retináról, megoldatlan a szakemberek munkájának minőségellenőrzése és a betegség progressiójának, illetve esetenként a terápia hatásosságának nyomon követése. A direkt tükrözéses módszer szenzitivitása olyan alacsony, hogy szűrőprogramok céljára nem alkalmas.

### **Az automatizált szűrést célzó eljárások**

A munka során i.) a könny proteomikai alapú módszertanokat illetve ii.) a képfeldolgozás, azon belül a MA-k detektálás alapú módszertanokat vizsgáltuk.

Ismert tény, hogy RD fennállása esetén a szemben lévő erek falának permeabilitása megváltozik, ezért azt feltételeztük, hogy ez hatással van a **könny fehérjeösszetételére** is. Számos közlemény írja le a fehérjeösszetétel megváltozását például a sebgyógyulás, illetve gyulladással megbetegedések kapcsán. Elektroforézises és tömegspektrometriás technikákkal végzett vizsgálatok támasztják alá azt, hogy az egészséges és DM-ban szenvedő

személyek könny fehérje összetétele jellemző eltéréseket mutat. Az utóbbi néhány évben számos potenciális RD biomarkert azonosítottak immunhisztokémiai, qPCR alapú mRNS-expressziós vizsgálatok vagy individuális fehérje koncentráció mérés segítségével. Ilyenek voltak például a szérumban lévő retinol kötő protein 4, a plazma apelin, az üvegtestben lévő PDGF és VEGF, a könny lipopoliszacharid (LPS)-kötő protein és szolubilis CD14, melyeknek emelkedett szintjét mérték proliferatív RD-ban szenvedő betegeknél. A projektek az egyes fehérjék célzott analízisének alapultak, így nem adtak rálátást a fehérjeösszetétel globális mintázatának változásaira. Egyes kutatócsoportok vizsgálták a neovascularizált membrán és a nem vaszkuláris epiretinális membrán összetételét proliferatív RD-ban szenvedő betegeknél poolozott mintákon LC-MS/MS metódust használva. A vizsgálatok a periostin és a pigment epitheliumderived factor (PEDF) szintjét találták szignifikánsan magasabbnak a RD-ás betegeknél.

Munkacsoportunk volt az első, aki a RD hatására bekövetkező könny proteomikai mintázat változását vizsgálta gépi tanulás alapú módszertannal.

A **képfeldolgozás alapú módszerek** a szemfenékről készített digitális fotón számítógépes algoritmusok segítségével ismerik fel azokat az



elváltozásokat, amelyek a RD-ra jellemzőek. Az egyik ilyen szemfenéki elváltozás, amelynek detektálására a napjainkban elterjedt algoritmusok többsége épül a MA.A detektálás során az algoritmusokat alkalmazó szoftver azonosítja azokat a szemfenéki laesioakat, amelyek feltételezhetően MA-k. Ezt követően a képeket egészséges és beteg (RD és non-RD) csoportba sorolja.

A digitalizált képek feldolgozását ún. előfeldolgozó lépésekkel kezdjük. Egyes szemfenéki laesioak jobban láthatóak a vörös szín kiszűrésével készült képeken. A színcsatornák megválasztása mellett fontos előfeldolgozó lépés a szemfenéki képek normalizálása (standardizálása) azok méretére, intenzitás eloszlására, zajcsökkentő elmosására vonatkozóan. Ezt követi azon algoritmusok alkalmazása, amelyek az egészséges szemfenéki képeket (pl. érhálózat, látóidegfő), illetve a kóros képeket (leggyakrabban az apró, vörös vagy a világos (fehéres/sárgás)) ismerik fel. A retina egyes elváltozásainak azonosítására kidolgozott legmodernebb eljárások tanítóalgoritmusokra építkeznek. Mivel a MA könnyen leírható, viszonylag konzekvens megjelenésű és érzékeny, korai indikátora a RD kialakulásának, az itt bemutatásra kerülő módszer kizárólag a MA-k detektálásán alapszik.

Munkánk során a ***könnyproteomikai és képfeldolgozás alapú módszertanokat kombináltuk***, amelyhez hasonló vizsgálatokat más munkacsoportok eddig nem végeztek.

## **2. Célkitűzések**

A könnyfehérje koncentrációk globális mintázatának vizsgálata RD-ban. A módszertan potenciális teljesítményének és annak értékelése, hogy az elemzéshez felhasznált markerek számának csökkentése milyen hatással van a vizsgálat szenzitivitás és specificitás értékeire. A könnyproteomikai és a képfeldolgozás alapú módszertanok egyéni eredményeinek javítása a potenciális kombinálhatóság kihasználásával (munka 2. fázisa) a RD diagnosztizálásában.

**Az 1. munkafázisában** a könnyben található fehérjék RD-markerként való alkalmazhatóságának vizsgálatára fókuszáltunk. Előkísérleteink során sikerrel azonosítottunk olyan fehérjét, melyek a RD-ban bekövetkező jellemző koncentrációváltozásuknak köszönhetően markerként használhatóak. Munkánk célkitűzései a következők voltak:

**Célkitűzés I.:** A könny fehérjék koncentrációjának egyedi vizsgálatát követően nagyszámú fehérje koncentrációjának egyidejű meghatározásával olyan globális mintázatokat kívántunk azonosítani, melyek a RD-ra jellemzőek. Feltételezésünk az volt, hogy ha egyidejűleg több fehérje RD okozta változását is figyelembe vehetjük, a szűrt populáció precízebb osztályozására leszünk képesek.

**Célkitűzés II.:** A könnyfehérje koncentrációk globális mintázatának vizsgálatát követően szerettük volna a vizsgált fehérjék számát csökkenteni úgy, hogy azzal a módszer diagnosztikai pontosságát ne csökkentjük jelentősen. Az alacsony prediktív értékkel bíró markerek kizárása, vagyis a vizsgált markerek számának csökkentése, az eljárás költséghatékonyságát növelni, kivitelezhetőségét pedig egyszerűbbé, gyorsabbá tenni volt hivatott. Elképzelésünk szerint ezzel az eljárás klinikai rutinban való alkalmazásának esélyét növelhetjük. Távlati célunk olyan a járóbetegellátásban is használható gyorseszteszt kidolgozása, mely a könnyből képes rövid idő alatt a RD jelenlétét megállapítani.

**A 2. munkafázis:** a könnyproteomikai és a képfeldolgozási módszertanok kombinálásán alapuló RD szűrő eljárás kidolgozása és vizsgálata. Ennek kapcsán az alábbi célkitűzést fogalmaztuk meg:

**Célkitűzés III.:**A szakirodalomban leírt eredményekkel való összehasonlíthatóság érdekében, célunk volt továbbá, hogy egy a mások által kidolgozott és publikált képfeldolgozás alapú eljárást implementáljunk. Az implementált módszer eljárás a RD korai jeleként számon tartott MA-k detektálásán alapult. Az eljárással kapott eredményeket reprodukálását követően tanulmányozni kívántuk, hogy a korábbiakban ismertetett könnyproteomikai eljárással kombinálva a kombinált módszertan teljesítménye meghaladja-e i.) a tisztán könnyproteomikai markerek használatára épülő eljárás teljesítményét ii.) és a tisztán képfeldolgozás alapú módszertan eredményét.

### **3. Módszerek**

#### **1. munkafázis – könnyproteomikai analízis**

**Betegvizsgálat:** 119 diabeteses beteg került bevonásra a Debreceni Egyetem Szemklinikájának járóbeteg szakrendeléséről. 73 betegnél csak az

egyik szemből vettünk könnymintát, a kizárási kritériumok teljesülése miatt, melyek az alábbiak voltak: szemfelszíni betegség (pl.: keratoconjunctivitis sicca, conjunctivitis, stb.), nem diabéteses eredetű szemfenéki elváltozások, operált szem, illetve a beteg együttműködési hajlandóságának hiánya. A RD diagnózisát 7-mezős fundus fotográfiák két független szemész általi vizsgálatával kaptuk. A könny minták gyűjtését arra kiképzett asszisztens végezte standardizált protokollt követve.

A **könnyproteomikai vizsgálatok** a Debreceni Egyetem, ÁOK, Biokémiai és Molekuláris Biológiai Intézetében történtek. A könny minták vizsgálatához modern nano-HPLC-vel kiegészített ESI-MS/MS tömegspektrométert használtunk. A minták globális fehérje koncentráció mintázatát vizsgáltuk, ami a könnyben lévő 34 fehérje koncentrációját és azok egymáshoz képesti arányát jelentette.

A **gépi tanuló eljárást** alkalmazó informatikai rendszerek tanuló adathalmaz használatával képesek tanulni, majd az új esetek inkomplett adatainak értékelésével megjósolni jövőbeni események bekövetkezését. Jelen esetben az eljárás inputját a könnyfolyadékban mért fehérje koncentrációk értékei jelentették. A tanuló fázisban a rendszer számára

megadtuk az esetek kimenetét is, ami jelen esetben a betegek RD státusa (RD/non-RD). A rendszer megtanulta, hogy mely proteomikai koncentráció mintázat jellemzi a RD illetve a non-RD eseteket. Az új eseteknél csak a könnyproteomikai adatok álltak a szoftver rendelkezésére, amely a tanult mintázatok alapján sorolja be az eseteket RD, illetve non-RD kategóriába.

A projekt során célunk volt különböző osztályozó eljárások teljesítményének értékelése, így ugyanazon az adathalmazon több vizsgálatot is végeztünk különböző osztályozó módszertanokat használva. Hat különböző modell/algorithmus teljesítményét hasonlítottunk össze SupportVectorMachine (SVM), RecursivePartitioning (rpart), Random Forest (randomForest), NaiveBayes (naiveBayes), LogisticRegression (logReg), K-NearestNeighbor (k-NN).Az elemzés során az "R" statisztikai keretrendszert használtuk. Az osztályozók a vizsgált adathalmazon elért teljesítményének monitorozásához k-szoros keresztvalidációs eljárást alkalmaztunk. A keresztvalidációs eljárás során az adathalmazt k-számú egyenlő részre osztottuk. Az első k-1 halmazt (tréning adathalmaz) használtuk fel a modellalkotáshoz, melyet aztán a k-adik halmazon tesztelünk (teszt halmaz). Az iterációk során ezt követően újabb k-1 darab részhalmazt alkalmaztunk a

rendszer tanítására, majd a maradék k-adik részhalmazon validáltuk a modellt.

Az első vizsgálati elrendezésben 34 fehérje koncentrációját használtuk bemenetként. Ezt követően a második vizsgálati elrendezésben kipróbáltuk, hogy mi történik abban az esetben, amikor a 34 fehérje koncentráció értéke közül csak azon fehérjék koncentráció értékeit használjuk, amelyek az előkísérletek során klasszikus statisztikai módszerek alkalmazásával biomarkerként alkalmazhatónak mutatkoztak szűrőeljárásokban való felhasználásra. Ekkor 6 fehérje koncentrációját adtuk meg a rendszer bemeneteként. A harmadik vizsgálati elrendezésben főkomponens analízissel (PCA) a 34 fehérjéből kiindulva próbáltuk csökkenteni az input változók, vagyis az elemzésbe bevont fehérjék számát.

## **2. munkafázis – a kombinált, képfeldolgozáson és könny proteomikai analízisen alapuló eljárás vizsgálata**

52 DM-ban szenvedő beteget vontunk be a vizsgálatok második fázisába. A betegvizsgálatok a fent leírtakkal egyező módon történtek. A proteomikai elemzés során ez esetben a 34 fehérje koncentráció értékeit használtuk. A különbséget az jelentette, hogy az input adathalmaz a

könnyproteomikai kísérletekből származó fehérje koncentráció értékek mellett a fundus fotók képfeldolgozási módszertannal történő vizsgálatának eredményeit, az azonosított MA-k számát is tartalmazta.

A vizsgálatokat három elrendezésben végeztük el: i.) első esetben kizárólag a könnyproteomikai; ii.) második esetben kizárólag a képi adatokat használtuk; iii.) a harmadik esetben mindkét adatféleséget használtuk. Ezt követően a három vizsgálati elrendezésben mértük az eljárás teljesítményét.

A **képfeldolgozási módszertan** alkalmazása során első lépésben egy MA detektort hoztunk létre, mely szintén gépi tanulás alapú módszertant alkalmazott a fundusfotókon meglévő MA-k felismerésére. Mivel a fundus képeken a zöld színcsatorna hordozza a legtöbb információt az érhálózatra vonatkozóan, ezért a háttérkorrekció során ebből került kivonásra a medián filterezett kép. A háttér korrigált képen kontraszt-limitált adaptív hisztogram kiegyenlítést (CLAHE) hajtottunk végre, melynek célja a képen lévő képletek és elváltozások kontrasztjainak felerősítése volt. A képfeldolgozás következő lépése a top hat transzformáció - kisméretű objektumok kivonására használt képfeldolgozási metódus - elvégzése volt morfológiai rekonstrukcióval. A rekonstruált kép 10x10-es korong-formát használ a kisméretű körkörös



elváltozások, vagyis a MA jelöltek azonosítására. Gradientboostingmachine (GBM) eljárást használtunk a MA jelöltek osztályozására a fenti jellemzők figyelembevételével. A betanításhoz olyan fundus fotókat használtunk, ahol a képen látható MA-k manuálisan be voltak jelölve. A következő lépésben ugyanilyen fotókat használtunk a MA detektor teljesítményének méréséhez.

A projekt során minden esetben a legjobban teljesítő osztályozó eljárást alkalmaztuk. A képfeldolgozáson alapuló modell esetén naiveBayes osztályozót, a könnyűprototípikus adatokat használó illetve a kombinált modell esetén pedig GBM-típusú osztályozó eljárást. A teljesítmény méréséhez ez esetben 10-szeres keresztvalidációs eljárást (10-szer ismételve) használtuk a fent leírt módon.

A gépi tanuló eljárás az alábbiak szerint használtuk (a kombinált módszertan példáján bemutatva): **Tanuló fázis:** Az esetek (szem) halmazából véletlenszerűen választottunk ki egy szubpopulációt, amit tanuló halmaznak neveztünk, majd az ismert diagnózisok alapján a szubpopuláció eseteit két csoportra osztottuk RD-re ill. non-RD-re. A klinikai diagnózis (RD/non-RD), a retina képeken azonosított MA-k száma, ill. a fehérje koncentráció értékek jelentették a gépi tanuló algoritmus inputját. Az algoritmus képes

megtanulni, hogy mely mintázatok a leginkább jellemzőek a RD ill. a non-RD csoportra. **Tesztelési fázis:** A következő lépésben a másik szubpopuláció (maradék esetek), a teszt halmaz adatait használtuk. Ebben a lépésben csak a fehérje koncentrációk értékeit, illetve a MA-k számát tekintettük az algoritmus inputjaként. A szem RD státuszát ebben a lépésben nem adtuk meg bemenetként. A tanuló algoritmus összevetette az új input jellemző mintázatait azokkal a RD, ill. non-RD esetekre jellemző mintázatokkal, amelyeket a tanuló fázisban megtanult. Ez alapján az algoritmus döntést hoz a teszt halmaz eseteire vonatkozóan. A modell teljesítményének értékeléséhez ezt az output-ot összevetjük az ismert klinikai diagnózissal.

## **4. Eredmények**

### **Az 1. munkafázis eredményei – könnyproteomikai analízis**

A hat különböző gépi tanuló eljárás közül az rpart bizonyult a leghatékonyabbnak, ha kizárólag a klasszikus statisztikai eljárásokkal azonosított hat marker fehérje adatait használtuk. Az értékelési mérőszámok azt mutatják, hogy az rpart osztályozó használatán alapuló modell teljesítménye felülmúlja a másik öt modell teljesítményét (0,74 szenzitivitás, 0,48 specificitás és 0,65 pontosság). Ugyan a Naiv Bayes osztályozó érte el a

legmagasabb szenzitivitás értéket (0,80), de a specificitása rendkívül alacsony (0,38). A Random Forest osztályozón alapuló modell kissé rosszabbul teljesített, mint a rpart(0,69 szenzitivitás, 0,45 specificitás és 0,65 pontosság).

A globális mintázat figyelembevételével történő vizsgálódás motivációját az adta, hogy feltételeztük, a markerként nem azonosított fehérjék is hordoznak olyan információt, amelynek felhasználása javíthatja a csak a hat marker használatával elérhető teljesítményt. A PCA alkalmazásának, vagyis a tanításhoz felhasznált adatmennyiség (input) csökkentése mögött az a törekvés állt, hogy egy a rutinban is alkalmazható, a teljes proteomikai panel tömegspektrométerrel való mérésénél egyszerűbben és olcsóbban megvalósítható eljárást dolgozzunk ki. Emellett szerettük volna megőrizni az eredeti eljárás teljesítmény értékeit. A vizsgálat eredményeként azt állapítottuk meg, hogy sem a 6 marker fehérjén, sem a PCA alkalmazásával elvégzett vizsgálat esetén nem kaptunk lényegesen eltérő eredményeket, mint amit a teljes proteomikai panel vizsgálatával kaptunk. A PCA során kapott első két komponens hordozta a teljes fehérje panel koncentráció értékei varianciájának 22%-át.

## **2. munkafázis eredményei – a kombinált eljárás vizsgálata**

A kombinált módszer alkalmazásával azt kívántuk bizonyítani, hogy két különböző típusú adatforrásból érkező adatféséség kombinált használatával az általunk felépített RD szűrésre szolgáló módszertan jobb teljesítményre képes, mint amire az egyes adatfésések felhasználásával felépített modellek külön-külön képesek.

Az **MA detektor önállóan** 0,84-es szenzitivitás illetve 0,81-es specificitás értéket ért el. A második modell, amely **kizárólag proteomikai** adatokra támaszkodott 0,87-es szenzitivitás értéket ért el 0,68-as specificitás mellett. A **kombinált eljárás** teljesítménye meghaladva mindkét vizsgált eljárás teljesítményét 0,93-as szenzitivitás és 0,78-as specificitás értéket ért el. Ez feltételezhetően annak köszönhető, hogy a két eltérő adatféséség egymástól független, de kiegészítő információt biztosított a kimenetekről az osztályozáshoz.

## 5. Megbeszélés

A munka célkitűzése egy olyan eljárás kidolgozása volt, amely alkalmas a diabeteses betegek, mint a RD kockázatának kitett populáció, elő-szűrésének végrehajtására. A DM-ban szenvedő, de RD képét nem mutató páciensek kiválogatásával csökken az a betegszám, amelyet a szűrőcentrumokban

dolgozó képelemzőknek kell megvizsgálnia. Mivel a jelenlegi szűrőeljárások igen munkaerő-igényesek és így költségesek, a szűrni kívánt populáció csökkentése jelentős költségmegtakarítással járhat.

**Akönnyminta** vétel nem invazív és viszonylag könnyedén végrehajtható, akár a házi orvos rendelőjében, vagy akár a beteg otthonában is elvégezhető.

A vizsgálatok során a modell által elért teljesítmény értékek alapján azonban úgy ítéljük meg, hogy a módszer jelenlegi formájában még nem alkalmas előszűrés-ként való alkalmazásra a klinikai rutinban, azonban egyértelműen alkalmas lehet a képfeldolgozás alapú eljárások teljesítményének javítására. Mivel ez a vizsgálat az első a maga nemében, még számos lehetőség van a módszer javítására.

Gotman és munkatársai által végzett **képfeldolgozás** alapú vizsgálat célkitűzése a biztosan negatív, normál képek eltávolítása a képadatbázisból a humán vizsgálók által végzett szűrést megelőzően, csökkentve ezáltal a vizsgálandó képadatbázis méretét, így a szűrés humán erőforrás igényét. Céljuk a 100%-os szenzitivitás elérése volt, amely mint ahogy az előre látható volt, alacsony specificitás értékkel párosul. A módszertan elő-szűrő

eljárásként alkalmazva a képek 26,4-38,1%-át volt képes kiszűrni, ami a munkaterhelés 1/3-os csökkenését eredményezi.

A munkacsoportunk által létrehozott modellek alkalmasak arra, hogy a proteomikai adatokkal párhuzamosan egyéb adatfészeségeket is inputként használjanak a vizsgálatok során. Ilyenek lehetnek például az anamnézisben szereplő klinikai adatok, vagy képi adatok, mint ahogy azt a saját munkánk második fázisának leírásánál láthattuk. Ha emeljük az inputként felhasznált adatfészeségek számát az szignifikánsan emelheti az eljárás teljesítményét. A **kombinált eljárással** általunk, vizsgálati körülmények között elért 0,93-as szenzitivitás és 0,78-as specificitás közelíti a humán vizsgáló által végzett szűrés célértékeit (0,80-as szenzitivitás és 0,95-as specificitás).

## 6. Összefoglalás

A retinopathia diabetica (RD) a cukorbetegség egyik leggyakoribb szövődménye. A vakságban szenvedő személyek 5%-ánál tehető felelőssé az állapot létrejöttéért. Azoknál a 2-es típusú DM-ban szenvedő betegeknél, akiknél a betegség 10 éve fennáll, 80% az esélye, hogy valamilyen fokú RDmár kialakult. Mivel a RD kezdetben tünetmentes, a diagnózis

időpontjában a betegség sokszor már igen előrehaladott stádiumban van. Időben történő diagnózissal és terápiával a RD progressziója jelentősen lassítható, így a DM-ban szenvedő populáció számára a rendszeres szemészeti szűrővizsgálaton (legalább évente) való részvétel ajánlott. A RD szűrés humánerőforrás igénye, így költségei is igen magasak, ami a szűrés automatizálásával csökkenthető lenne. Jelen munka célja a RD szűrésére alkalmas automatizált eljárás létrehozása.

Az első munkafázisban olyan módszertant dolgoztunk ki, amely a könnyű fehérjék RD okozta koncentráció változásainak vizsgálatán alapszik. A DM-os betegek osztályozásához (RD/non-RD) gépi tanuló eljárást használtunk, mely a fehérje koncentrációk globális mintázatának megváltozása alapján sorolta a betegeket egyik, vagy másik csoportba. A vizsgálatok eredményeként kapott viszonylag alacsony szenzitivitás (0,74) és specificitás (0,48) értékek alapján arra a következtetésre jutottunk, hogy a módszer jelenlegi állapotában önállóan még nem érett a klinikai rutinban való alkalmazásra. Azonban más típusú eljárással kombináltan használva alkalmas lehet a másik módszer teljesítményének javítására.

A második munkafázisban a könny proteomika alapú módszertant digitalizált retina képek automatizált feldolgozásával kombináltuk. A vizsgálatok során a képfeldolgozás alapú módszer önállóan 0,84-es szenzitivitás és 0,81-es specificitást ért el. A könnyproteomikai módszertan ebben az esetben 0.87-es szenzitivitás és 0,68-as specificitást produkált. A kombinált eljárás inputját a digitális fundus fotón azonosított MA-k száma és a könnyfehérje koncentrációs értékek jelentették. Az eljárás 0.93/0,78-os szenzitivitás/specificitás értéket ért el. Mivel a két adatféleség független, de egymást kiegészítő információt hordozott a vizsgálat várható kimenetére vonatkozóan, a három módszertanból ez bizonyult a legeredményesebbnek közelítve azt a minimum értéket, melyet a British Diabetic Association a szűrőprogramok számára meghatározott (legalább 0.80 szenzitivitás és 0.95 specificitás).

## **7. Új eredmények**

### **1. munkafázis:**

- (1.) A különböző könny fehérjék koncentrációinak globális mintázatát gépi tanulás alapú módszerekkel elemezve, sikerrel azonosítottunk



olyan mintázatokat, amelyek a RD-ban szenvedő betegekre jellemzőek voltak.

- (2.) Olyan eljárást alkottunk, amely a könnyfehérje koncentrációk értékeit figyelembe véve, (kísérleti modellben) a DM-ban szenvedő betegeket RD és non-RD kategóriákba volt képes osztályozni.
- (3.) Sikerral hoztunk létre olyan szűrőeljárást, amely az előkísérletek során a kutatócsoport által klasszikus statisztikai módszerekkel azonosított 6 markerfehérje koncentrációinak egyidejű használatával képes a szűrt populáció DM-ban szenvedő betegeinek osztályozására (RD, illetve non-RD csoportokba).

## **2. munkafázis:**

- (4.) A területen elsőként sikerrel kombináltunk képfeldolgozás és könny fehérje proteomika alapú eljárást a RD szűrésére (kísérleti modellen).
- (5.) A kombinált eljárás teljesítménye meghaladta az egyedi (képfeldolgozás, ill. könnyproteomikai eljárás) eljárások teljesítmény mutatóit, vagyis a független, eltérő forrásból származó információk

együttes használata egy az egyedi eljárásoknál erősebb osztályozót eredményezett.

## **TÁMOGATÁS**

A munka sikeres megvalósításához az alábbi pályázat jelentős mértékben hozzájárult: Vaszkuláris és kardiális kutatóhálózat: Az ér- és a kardiovaszkuláris betegségek pathomechanizmusa, diagnosztikai, farmakológiai befolyásolhatóságuk az alapkutatás szintjén”. TAMOP-4.2.2.A-11/1/KONV-2012-0045.



DEBRECENI EGYETEM  
EGYETEMI ÉS NEMZETI KÖNYVTÁR  
PUBLIKÁCIÓK



Iktatószám: DEENKÉTK/399/2014.  
Tételszám:  
Tárgy: PhD Publikációs Lista

Jelölt: Török Zsolt  
Neptun kód: LQ4VE7  
Doktori Iskola: Klinikai Orvostudományok Doktori Iskola  
MTMT azonosító: 10046410

A PhD értekezés alapjául szolgáló közlemények

1. **Török, Z.**, Pető, T., Csősz, É., Tukacs, E., Molnár, Á.M., Berta, A., Tőzsér, J., Hajdu, A., Nagy, V., Domokos, B., Csutak, A.: Combined methods for diabetic retinopathy screening, using retina photographs and tear fluid proteomics biomarkers.  
*J. Diabetes Res.* 2014, 8 p., 2014.  
DOI: <http://dx.doi.org/http://dx.doi.org/10.1155/2014/623619>  
IF:3.536 (2013)
2. **Török, Z.**, Pető, T., Csősz, É., Tukacs, E., Molnár, Á., Maros-Szabó, Z., Berta, A., Tőzsér, J., Hajdu, A., Nagy, V., Domokos, B., Csutak, A.: Tear fluid proteomics multimarkers for diabetic retinopathy screening.  
*BMC Ophthalmol.* 13 (1), 8 p., 2013.  
DOI: <http://dx.doi.org/10.1186/1471-2415-13-40>  
IF:1.075
3. Csősz, É., Boross, P., Csutak, A., Berta, A., Tóth, F., Póliska, S., **Török, Z.**, Tőzsér, J.: Quantitative analysis of proteins in the tear fluid of patients with diabetic retinopathy.  
*Journal of Proteomics.* 75 (7), 2196-2204, 2012.  
DOI: <http://dx.doi.org/10.1016/j.jprot.2012.01.019>  
IF:4.088





### További Közlemények

4. Czimmerer, Z., Hulvely, J., Simándi, Z., Varallyay, É., Havelda, Z., Szabó, E., Varga, A., Dezső, B., Balogh, M., Horváth, A., Domokos, B., **Török, Z.**, Nagy, L., Bálint, B.L.: A versatile method to design stem-loop primer-based quantitative PCR assays for detecting small regulatory RNA molecules.  
*PLoS One*. 8 (1), 10 p., 2013.  
IF:3.534
5. Laczik, M., Tukacs, E., Uzonyi, B., Domokos, B., Doma, Z., Kiss, M., Horváth, A., Batta, Z., Maros-Szabó, Z., **Török, Z.**: Geno viewer, a SAM/BAM viewer tool.  
*Bioinformatics*. 8 (2), 107-109, 2012.
6. Gó, B., Balla, A., Tukacs, E., Nagy, I., **Török, Z.**: A Novel High Accuracy Algorithm for Reference Assembly in Colour Space.  
*Int. J. Biometr. Bioinform*. 6 (4), 92-104, 2012.
7. Tukacs, E., Korotij, Á., Maros-Szabó, Z., Molnár, Á.M., Hajdu, A., **Török, Z.**: Model requirements for Biobank Software Systems.  
*Bioinformatics*. 8 (6), 290-292, 2012.  
DOI: <http://dx.doi.org/10.6026/97320630008290>
8. Antal, B., Hajdu, A., Maros-Szabó, Z., **Török, Z.**, Csutak, A., Pető, T.: A two-phase decision support framework for the automatic screening of digital fundus images.  
*J. Comput. Sci.* 3 (5), 262-268, 2012.  
DOI: <http://dx.doi.org/10.1016/j.jocs.2012.01.001>
9. Inczedy-Farkas G., Benkovits J., Balogh N., Álmos P., Scholtz B., Zahuczky G., **Török Z.**, Nagy K., Réthelyi J., Makkos Z., Kassai-Farkas Á., Égerházy A., Tüzkő J., Janka Z., Bitter I., Németh G., Nagy L., Molnár M.J.: Magyar szkizofrénia-biobank a szkizofrénia kutatás és a személyre szabott orvoslás szolgálatában.  
*Orv. Hetil.* 151 (35), 1403-1408, 2010.  
DOI: <http://dx.doi.org/10.1556/OH.2010.28943>





**DEBRECENI EGYETEM**  
**EGYETEMI ÉS NEMZETI KÖNYVTÁR**  
**PUBLIKÁCIÓK**



10. Hunyadi J., Bihari I., Nagy I., Török Z., Hamar P.: Krónikus sebek kezelése a gyógyulás különböző fázisaiban.  
*Értekezések. 11 (2), 57-65, 2004.*

**A közlő folyóiratok összesített impakt faktora: 12,233**

**A közlő folyóiratok összesített impakt faktora (az értekezés alapjául szolgáló közleményekre): 8,699**

A DEENK a Jelölt által az iDEa Tudóstérbe feltöltött adatok bibliográfiai és tudományometriai ellenőrzését a tudományos adatbázisok és a Journal Citation Reports Impact Factor lista alapján elvégezte.

Debrecen, 2014.12.04.

