

A világháló nyújtotta elemző eszközök az emlő rosszindulatú daganatainak vizsgálatában

NAGY ÁDÁM^{1,2}, GYÖRFFY BALÁZS^{1,2}

¹MTA TTK Lendület Onkológiai Biomarker Kutatócsoport, ²Semmelweis Egyetem II. Sz. Gyermekklinika, Budapest

Levelezési cím:

Nagy Ádám, MTA TTK Lendület Onkológiai Biomarker Kutatócsoport, H-1117 Budapest, Magyar Tudósok körútja 2., tel.: +36 1 382 6745, e-mail: nagy.adam@ttk.mta.hu

Közlésre érkezett:

2015. november 20.

Elfogadva:

2015. december 18.

Az internet az emlő rosszindulatú daganatainak vizsgálatában alkalmazható eszközök egy új generációját hozta el az elmúlt évtized folyamán. Ezek között a legnagyobb csoportban levő programok a várható túlélés, illetve a terápiás válasz előrejelzését klinikai adatok alapján végzik el. Ide tartozik az Adjuvant! Online, az MSKCC és az MD Anderson nomogramjai, valamint az angliai PREDICT honlapja. Ezek mindegyikében közös, hogy a beteg megadott paramétereit alapján kikeresi, hogy mi történt a beteghez hasonló betegekkel korábban. Új diagnosztikus biomarkerek azonosításának forrása lehet génexpresszió vagy mutáció – meglevő, teljesgenom-szintű adatokat tartalmazó adatbázisok kiaknázásán alapszik az egyes gének vagy génmintázatok prognosztikus értékét számoló KMplot.com. Az adott gének mutációs adatai és a várható túlélés közötti kapcsolatot lehet a cBioportalban, illetve a G-2-O-ban vizsgálni. Géncsipekkel lement genomszintű génexpressziós adatok alapján ad diagnózist a RecurrenceOnline program. Az emlőtumor kialakulásának valószínűségét egy poligénes modell alapján, klinikai adatok felhasználásával becsüli meg a BOADICEA. Az emlőtumorok vizsgálatával foglalkozó onkológusok, patológusok, valamint kutatók számára készült összefoglaló közleményünkben ezen, valamint további elemző rendszerek célját, alkalmazási területeit és a rendszerek mögötti adatbázisokat foglaljuk össze. Magyar Onkológia 60:273–280, 2016

Kulcsszavak: emlőtumor, biomarker, prognózis, génexpresszió, mutáció

A new generation of internet-based diagnostic and research tools have arrived in the last decade. The most extensive group of these includes programs predicting the expected survival mainly by utilizing clinical data of the patient. This includes Adjuvant! Online, the MSKCC and MD Anderson nomograms and the UK-based PREDICT algorithm. A common feature of all these is the comparison of the given patient to previously treated breast cancer samples, and evaluating the clinical outcome of these previous patients. New diagnostic biomarkers can be gene expression or mutation based. Of these, large transcriptomic databases lay the basis for the KMplot.com analysis platform which is capable to assess the prognostic value of a selected gene or gene set. The link between a given mutation and survival is the focus of the cBioportal and the G-2-O software. Diagnosis is based on a transcriptome-level data derived using gene chips in the RecurrenceOnline algorithm. A risk of breast cancer development is assessed by a polygenic model in BOADICEA. In our review we target oncologists, pathologists and breast cancer researchers and provide a comprehensive summary of these and other analysis platforms.

Nagy Á, Györfly B. Internet-based opportunities in breast cancer diagnostics and research. Hungarian Oncology 60:273–280, 2016

Keywords: breast cancer, biomarker, prognosis, gene expression, mutation

BEVEZETÉS

Az emlő rosszindulatú daganatai vezetnek a tumoros betegségek előfordulását a nők körében, a daganattal kapcsolatos halálesetek szempontjából pedig a második helyen vannak a ranglétrán (1). A terápia során a gyógyszerre adott választ és a kezelés végkimenetelét a daganat számos klinikai és patológiai tulajdonsága meghatározza, mint a tumor mérete, szövettani differenciáltsága, nyirokcsomóstátusza, az ösztrogénreceptor (ER) és a progeszteronreceptor (PR), a humán epidermális növekedési faktor receptor 2 (HER2) és az MKI67 proliferációs marker kifejeződése (2). Az elmúlt másfél évtized folyamán géncsipadatok alapján az emlőtumort klinikailag négy fő altípusra osztották, amelyek a hagyományos immunhisztokémia alapú markerekkel is korrelálthatóak – ezek a luminális-A (ER+/HER2-/alacsony MKI67), luminális-B (ER+/HER2+ és ER+/HER2-/magas MKI67), HER2 (ER-/HER2+) és bazális tumorok (ER-, PR-, HER2-) (3, 4).

A luminális emlőtumorok (luminális-A és luminális-B), mint hormonreceptor-pozitív daganatok fontos sajátossága az ösztrogénreceptor gén és/vagy a progeszteronreceptor gén kifejezése, melyek miatt e tumorok endokrin terápiás kezelése lehetséges. Fontos megemlíteni azonban, hogy a luminális-B alcsoport ~30%-a a HER2 receptort is kifejezi (5). Ezen altípusba tartoznak azok a tumorok is, amelyekben a sejtproliferációhoz kapcsolódó gének – CCNB1, MKI67, MYBL2 – magasabb expressziót mutatnak. A luminális-B altípushoz rosszabb prognózis társul, az endokrin kezelés kevésbé hatásos és a daganat kiújulásának a kockázata is magasabb (6). E HER2-pozitív tumorok esetén az endokrin terápiát célzott HER2-ellenes terápiával kombinálva a daganat kiújulásának a kockázata csökkenthető (7).

A HER2-pozitív emlőtumorok legfontosabb tulajdonsága a HER2 gén amplifikációja vagy fokozott expressziója (3). Előfordulásuk főleg a menopauza után lévő nők körében gyakori, és a betegség prognózisa a luminális emlőtumorokéhoz képest rosszabb (4, 8). E tumortípus esetében alkalmazott célzott terápia alapja a HER2-ellenes monoklonális antitestek (pl. trastuzumab, pertuzumab) alkalmazása kemoterápiával kombinációban (7).

A bazális emlődaganatok fontos jellemzője, hogy nem expresszálják sem ösztrogén- vagy progeszteronreceptort, sem HER2-fehérjét (3). Génexpressziós vizsgálatok alapján a bazális emlőtumorok megkülönböztető tulajdonsága a citokeratin 5, 6 vagy 17 expressziója is, amely egyébként a bazális epithelsejtekben is megfigyelhető (3, 4). Bár a bazális emlőtumoros betegek műtétet megelőző ún. neoadjuváns kemoterápiára érzékenyek (9), azonban ennek ellenére is a bazális emlődaganatok rossz prognózissal társulnak, és az érintett betegek halálozási rátája a többi altípushoz képest magasabb (4).

A fentebb említett négy altípuson kívül létezik egy ötödik altípus, amely hasonlóan a bazális emlőtumorokhoz receptor-negatív, azonban fontos elkülönítő bélyege, hogy a daganatsejtek közötti „tight junction” (szoros sejt közötti kapcsolatok) kialakításában részt vevő fehérjék – claudin 3, 4, 7 és az

E-cadherin – alacsony expressziót mutatnak, így ezen altípus a „claudin-low” elnevezést kapta (10, 11). Hasonlóan a bazális emlőtumorhoz, a „claudin-low” alcsoport egyes kemoterápiás szerekekkel szemben érzékeny, azonban ennek ellenére a betegek relapszusmentes, valamint teljes túlélése alacsony (12).

Az egyre pontosabb patológiai és molekuláris osztályozásnak, a hormon- és célzott terápiának köszönhetően az emlő rosszindulatú daganatainak kezelése szinte már sikertörténet is lehetne, hiszen 1990 óta a mortalitás majdnem a felére csökkent az Amerikai Rákkutató Társaság epidemiológiai adatai alapján. Ennek ellenére a betegek igen jelentős részében továbbra is relapszus alakul ki. Új diagnosztikai markerek azonosítására lesz szükség ahhoz, hogy e minták esetében is ki tudjuk választani a legmegfelelőbb kezelést. Az elmúlt néhány évben egy új, szoftver alapú diagnosztikai paletta vált elérhetővé az emlődaganatok hatékonyabb osztályozására, valamint további biomarkerek azonosítására. Összefoglaló referátumunkban az e körbe tartozó, a világhálón keresztül elérhető eszközöket mutatjuk be.

KLINIKAI ADATOK FELDOLGOZÁSA

Több interneten elérhető alkalmazás létezik, amelyek a beteg klinikai adatait felhasználva nyújtanak információt a betegség várható előrehaladásáról, a terápiás lehetőségekről, többek között az adjuváns kezelések alkalmazásának várható sikerességéről. Ezek közül az Adjuvant! Online az egyik legismertebb, az interneten szabadon hozzáférhető program, melynek segítségével meghatározható az adjuváns kezelések várható hatásossága (13).

Az Adjuvant! Online (<https://www.adjuvantonline.com/>) alapját a SEER (Surveillance, Epidemiology, and End-Results) program adatai képezik, amely több mint harmincezer, az Amerikai Egyesült Államokban kezelt emlődaganatos nő 10 éves teljes túlélési adatát tartalmazza (14). Az Adjuvant! Online a betegek klinikai adatait használja fel, úgy mint a beteg életkorát, nemét, társuló betegségeket, a tumor méretét, nyirokcsomóstátusát, ER-státusát és a szöveti differenciáltság mértékét. Segítségével akár 10 évre előre, egyénre szabottan számítható az adjuváns terápia alkalmazása vagy mellőzése mellett az emlőtumor várható kiújulása és a daganattal kapcsolatos halálozás valószínűsége. Az előrejelzés során külön megadja, hogy hormonterápia, kemoterápia vagy kombinációs terápia alkalmazása várhatóan milyen eredményt fog hozni. A könnyű használhatóság is segítette a rendszer elterjedését, a gyakori használatot az első közleményre érkezett több mint 700 idézés is alátámasztja. Az Adjuvant! Online hátránya azonban, hogy nem veszi figyelembe a daganat HER2-, PR-, illetve proliferációs státuszát, amelyek fontos prognosztikus és prediktív markerek, így az emlődaganat prognózisának meghatározásában elérhető pontossága korlátozott (13). Az Adjuvant! program által elvégzett számítás eredményét az 1. ábrán¹ mutatjuk be.

¹Hozzáférés időpontja: 2015. IX. 30.

ADJUVANT! ONLINE

Decision: No additional therapy



Decision: Hormonal therapy



Decision: Chemotherapy



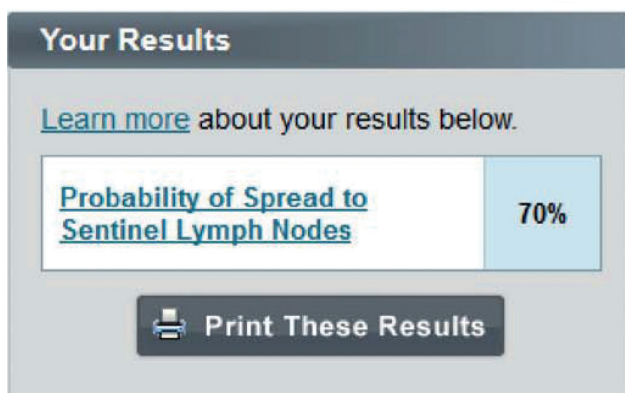
Decision: Combined therapy



1. **ÁBRA.** Az adjuváns kezelések várható sikerességének meghatározására alkalmas Adjuvant! Online

A második online eszköz az MSKCC (Memorial Sloan Kettering Cancer Center) által létrehozott három prognosztikus nomogramot tartalmazza (<http://nomograms.mskcc.org/breast/>), amelyek a klinikai döntéshozatalt támogatják a legmegfelelőbb terápiás kezelés kiválasztásában. A „sentinel lymph node” nomogrammal megbecsülhető, hogy a betegekben milyen valószínűséggel terjed az emlődaganat az őrszem (sentinel) nyirokcsomókba. A számítást számos klinikai paraméter, mint a beteg életkora, a tumor mérete, limfóvaszkuláris terjedési képessége, elhelyezkedése, multifokálisitása (több tumorgóc jelenléte), típusa, stádiuma, valamint a hormonreceptorok jelenléte alapján végzi (15). Az „additional node metastasis”

MSKCC



2. **ÁBRA.** Az MSKCC nomogram által készített előrejelzés

alkalmazás a tumor távolabbi nyirokcsomókba történő terjedésének valószínűségét számítja ki azon egyének esetében, akiknél a daganat az őrszem nyirokcsomókat már elérte. Az előző nomogramhoz szükséges adatokon kívül itt a pozitív és negatív őrszem nyirokcsomók száma és ezek detektálásának módja is szükséges (16). A „DCIS recurrence” alkalmazás segítségével megtudható, hogy egy DCIS emlőtumorok a műtétet követően mekkora a kiújulási valószínűsége az adott adjuváns kezelés eredményeképpen. A program a daganat diagnosztizálásának időpontja, családtagok érintettsége, a betegség észlelésének módja, adjuváns sugár-/adjuváns endokrin terápia megléte, a szöveti differenciáltság mértéke, nekrozis előfordulása és a műtét éve alapján határozza meg a daganat öt- és tízéves kiújulásának valószínűségét és ad ez alapján támpontot a leghatásosabb terápia kiválasztásához (17). A „DCIS recurrence” alkalmazás példa előrejelzését a 2. ábrán² mutatjuk be.

Szintén nomogram alapúak a Texasi Egyetemen az MD Anderson Cancer Center által kifejlesztett online rendszerek (<http://www.mdanderson.org/education-and-research/resources-for-professionals/clinical-tools-and-resources/clinical-calculators/index.html>). Itt is több nomogramot fejlesztettek ki az emlőtumorra specifikusan, amelyek előre jelezhetik a neoadjuváns kemoterápia hatásosságát, a sebészeti eljárás utáni túlélést, az őrszemnyirokcsomó-pozitív beteg neoadjuváns kemoterápiás kezelés után vagy kezelés nélküli prognózisát, valamint az őrszemnyirokcsomó-pozitív betegek távolabbi nyirokcsomóiban megjelenő áttétek valószínűségét. Ez utóbbi az MSKCC nomogramhoz hasonlóan képes kiszámítani, hogy az őrszem nyirokcsomókat elérő daganat esetén mennyi a valószínűsége a tumor távolabbi nyirokcsomókban való további áttétképzésének (18). A beálítható klinikai paraméterek közé tartozik a tumor mérete, típusa, a pozitív őrszem nyirokcsomók száma, a limfóvaszkuláris invázió, amelyek az MSKCC rendszerben is megtalálhatóak (16). Egyedi prognosztikus paramétereikhez sorolható az eltávolított nyirokcsomók száma, az áttét mérete, valamint a nyirokcsomón kívüli érintettség. Információt kaphatunk arról, hogy az adott beteg esetén a hónalji nyirokcsomók eltávolításának milyen hatása van, illetve van-e egyáltalán hatása a túlélésre (18).

A Cambridge-i egyetemen készült a PREDICT algoritmus (<http://www.predict.nhs.uk/predict.html>), amely 5694 angliai beteg adatai alapján szintén az öt- és tízéves várható túlélést jelzi előre kemoterápiával és kemoterápia nélkül (19). A PREDICT-et eredetileg ötezer beteg adatain tesztelték, azonban egy későbbi, elsősorban kanadai mintákat felhasználó vizsgálatban az Adjuvant! Online eredményeivel közvetlenül is összehasonlították. E vizsgálat eredményeképpen a két előrejelző hasonló hatásosságot ért el (20). A PREDICT továbbfejlesztett verziója már figyelembe veszi a tumor HER2- és MKI67-státuszát is, ami lehetővé teszi,

²Hozzáférés időpontja: 2015. IX. 23.



3. ÁBRA. A várható túlélés meghatározása a PREDICT algoritmussal

hogya a többi fenti rendszernél nagyobb hatásosságot érjen el a várható prognózis előrejelzésében (21). A PREDICT rendszerrel készített, a várható túlélést mutató példa eredménygrafikont a 3. ábrán³ mutatjuk be.

BIOMARKEREK AZONOSÍTÁSA ÉS IGAZOLÁSA

Számos gén expresszióbeli megváltozása összefüggést mutat a túléléssel és/vagy a terápiás válasszal. Egyes esetekben egy gén, azonban számos esetben több gén együttes működésbeli változása befolyásolhatja a tumor viselkedését (22). Az emlőtumor esetében számos prognosztikus és prediktív tulajdonságokkal bíró gént (biomarkert) azonosítottak, amelyek a daganatos betegek terápiájában jelentősen megkönnyíthetik a klinikai döntéshozatalt, így növelhetik a terápia hatékonyságát (23). Ezek közül kiemelkedő jelentőségű maga az ösztrogénreceptor, amely ma is az onkológiában alkalmazott célzott terápia zászlóshajója. A második sikertörténet a HER2 gén és célzott gátlása: a magas HER2-kifejeződésű, rossz túlélésű csoport várható kimenetelét a HER2-ellenes terápia alkalmazása jelentősen megváltoztatta. Mára a HER2-pozitív betegek várható prognózisa jobb, mint a HER2-negatív betegek

³Hozzáférés időpontja: 2015. IX. 25.

prognózisa (7). További biomarkerek azonosítására, illetve új biomarkerjelöltek igazolására több, online elérhető program is létezik, amelyekben közös, hogy korábban publikált minták klinikai, génexpressziós vagy mutációs adatait felhasználva tesznek lehetővé egy automatizált retrospektív elemzést.

A kutatócsoportunk által létrehozott „Kaplan–Meier-plotter” (<http://kmpplot.com/analysis/>) géncsípérések során meghatározott 54 675 gén expressziós adatait használva határozza meg a génnek a betegség prognózisában betöltött jelentőségét. A rendszer a Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>) és az ArrayExpress (<https://www.ebi.ac.uk/arrayexpress/>) adatbázisaiból összesen 5143 emlő- (24), 1065 gyomor- (25) és 2437 tüdőrák (26) mintát tartalmaz. A program lehetővé teszi egy vagy több gén kapcsolatának vizsgálatát a relapszusmentes, távoliáttét-mentes és teljes túléléssel. A program a kiválasztott gén alacsony vagy magas expressziója alapján a betegeket két csoportra osztja, majd az így létrehozott két csoportot hasonlítja össze Cox-regresszióval és készíti el a kapcsolódó Kaplan–Meier-görbét (24). Az elemzés során több minőség-ellenőrzési lépés is beépítésre került, mint pl. egy adott gén mérésére a legmegbízhatóbb próba kiválasztása (27), a többször leköszölt géncsipek kizárása (28), valamint a géncsipek minőségének ellenőrzésére javasolt nyolc minőségi paraméter alkalmazása (29). Az adatbázis felosztásával lehetőség van egyes leszűkített betegcsoportok kiválasztására, amelyekben az analízist el szeretnénk végezni, mint például az alkalmazott endokrin terápia alapján (30), vagy pedig a TP53 gén mutációs státusza alapján (31). Az algoritmus mRNS-ek mellett mikroRNS-ek prognosztikus értékét is tudja számolni (32). A rendszer használatát mutatja, hogy havonta több mint ötezer elemzést készítenek a felhasználók.

Az emlőtumor vizsgálatára alkalmas, az interneten szabadon hozzáférhető másik rendszer az MSKCC által létrehozott cBioportal (<http://www.cbioportal.org/>), amely komplex genomi adatok interpretálásával, feldolgozásával és az eredmények grafikus szemléltetésével új potenciális biomarkerek azonosítására képes. A cBioportal működése a TCGA (The Cancer Genome Atlas) adatainak integrálásán alapul (33). A cBioportal számos lehetőséget biztosít a daganatok szekvenciaalapú molekuláris jellegzetességeinek és folyamatainak tanulmányozására, többek között alkalmas egy vagy több gén mutációs mintázatának vizsgálatára, valamint a betegre jellemző genomi eltérések összesítésére. A cBioportalon belül található OncoPrint modul segítségével képet kaphatunk arról, hogy egyes tumortípusokra nézve a betegek hány százaléka tartalmaz a kiválasztott gént vagy génekre jellemző szomatikus mutációkat, valamint kópiaszám-változásokat. Kettő vagy annál több gén vizsgálata esetén az is meghatározható, hogy az egyes mutációk milyen valószínűséggel fordulnak elő együtt (33).

Szintén mutációs adatokat használ a „Genotype to outcome” honlap (<http://www.g-2-o.com/>). A G-2-O kon-

cepciója, hogy egy mutáció nem direkt önmagában fejt ki hatását, hanem egyszerre több további gén expresszióját is befolyásolja. Ahhoz, hogy a mutáció hatását meg tudjuk becsülni, e gének összességének hatását kell megvizsgálni. A G-2-O rendszernek azt kell megadni, hogy az adott betegben mely génekben volt mutáció, majd ezek alapján megkeresi a mutációt hordozó betegekben eltérő kifejeződést mutató géneket, végül pedig a Kaplan–Meier-rajzolásnál fentebb már említett adatbázis felhasználásával kiszámolja, hogy milyen lesz a várható túlélés [34]. A G-2-O algoritmus emellett lehetővé teszi azt is, hogy a túlélést befolyásoló, a tumor progressziójában szerepet játszó mutációkat el tudjuk különíteni a tumor progressziója folyamán létrejövő potyautas („passenger”) mutációktól, amint ezt igazoltuk is a KRAS onkogén esetén [35].

HÁLÓZATELEMZÉS

A rosszindulatú daganatok a komplex, többtényezős, úgynevezett multifaktoriális megbetegedések csoportjába tartoznak, amelyeket mind genetikai, mind környezeti faktorok befolyásolnak. A több mint két évtizeddel ezelőtt induló Humán Genom Projekt [36] célja volt a betegségek kialakulásáért felelős genetikai faktorok azonosítása. A teljes humán genom szekvenálásával, valamint az ehhez kapcsolódó további projektekkel – pl. a HapMap a genomszintű asszociációs vizsgálatokkal [37] – lehetővé tették további, a betegségek kialakulásával kapcsolatos gének azonosítását [38].

Az 1995-ben a Japán Humán Genom Program részeként létrehozott, a genomszekvencia-adatokat integráló Kyoto Encyclopedia of Genes and Genomes (<http://www.genome.jp/kegg/>) mára a poszt-genom analízis egyik népszerű online eszköze lett [39]. A KEGG egy integrált adatbázis, amely 16 részadatbázisból épül fel, tartalmazza a metabolikus és szabályozó útvonalakat, valamint a jelátviteli, sejtciklus-, illetve a fejlődésben szerepet játszó kaszkádok összességét. E részadatbázisok három főbb kategóriatípusra bonthatóak: genomi, kémiai és hálózatszintű információt tartalmazó adatbázisokra. A program minden útvonalat a KegSketch szoftverrel rajzol meg, és összegzi a hozzá kapcsolódó irodalmat [40].

A DAVID (Database for Annotation Visualization and Integrated Discovery) (<https://david.ncifcrf.gov/>) egy másik integrált adatbázis, amely nagyszámú gén/fehérje funkcionális vizsgálatára alkalmas [41]. A DAVID fő alkalmazásai között van a „Functional Annotation”, melynek segítségével a kiválasztott génekre a program úgynevezett annotációs csoportokat képez, és megmutatja a csoporton belüli hasonlóság mértékét [41]. A „Gene Functional Classification Tool” a megadott géneket funkciójuk alapján csoportosítja, így azon géneket, amelyek hasonló folyamatokban játszanak szerepet, a program egy géncsoportba sorolja [42].

Az elmúlt évtizedben az újgenerációs szekvenálási és a géncsip-technológiák fejlődése lehetővé tette a daganatra jellemző komplex genomi eltérések különböző szintjeinek egyidejű vizsgálatát. A Yale Egyetem, az MTA és a Sem-

melweis Egyetem együttműködésében létrehozott NetGO (<http://netgoplot.org/>) egy olyan online rendszer, amely az emlődaganatos betegek különböző adatainak – géncsip, génkópiaszám-eltérés, szomatikus mutációk – integrálásával lehetőséget nyújt az egy adott páciensre jellemző komplex genomi rendellenességek biológiai, metabolikus és celluláris folyamatokban betöltött szerepének vizsgálatára [43]. A rendszer a genomi rendellenességek és az egyes folyamatok közötti összefüggések vizsgálatához a Gene Ontology (GO) konzorcium (<http://geneontology.org/>) által létrehozott genomannotációs adatbázist használja fel, amely információt tartalmaz a gének molekuláris funkciójáról, biológiai folyamatokban betöltött szerepéről, valamint az egyes folyamatok közötti kapcsolatokról. A NetGO program arról ad információt, hogy az adott betegnél e folyamatok közül melyek játszanak szerepet.

DIAGNÓZIS EGYETLEN BETEG SZÁMÁRA

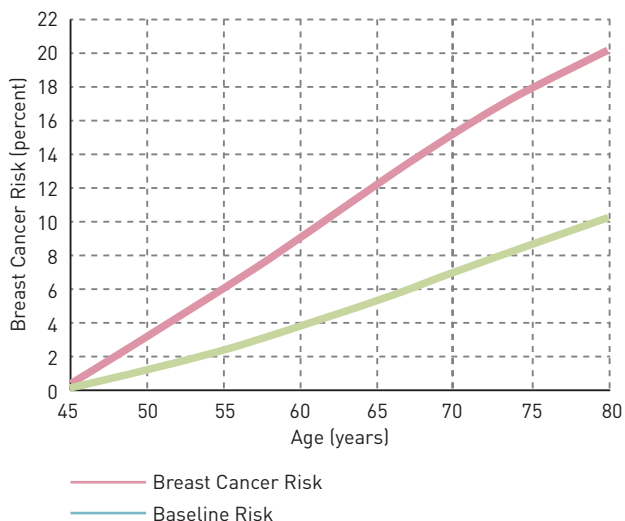
Az elmúlt évtizedben számos, génexpresszió alapuló egy- és többgén tesztet fejlesztettek ki az emlődaganatok pontosabb prediktív és prognosztikus diagnózisához. Ezek során a cél a beteg számára leghatékonyabb terápia alkalmazásán keresztül nemcsak a várható túlélés időtartamának növelése volt, de a szükségtelen terápia elhagyásán keresztül a költségek és mellékhatások csökkentése is [44].

A jelenleg a daganatok diagnosztizálására alkalmas, kereskedelmi forgalomban lévő tesztek fő hiányossága, hogy a hatásosságukat független vizsgálat nem igazolta. A legtöbb tesztet ráadásul meglehetősen kisszámú és viszonylag heterogén betegcsoporton fejlesztették ki, ezért az előrejelző képességük az e csoportokba nem illő betegek korlátozott [45]. Például a 70 gént vizsgáló MammaPrintet mindössze 78 betegminta alapján [46], a 21 gént vizsgáló OncotypeDx-et 233 ER-pozitív, nyirokcsomó-negatív betegminta alapján [47], a 93 gént vizsgáló Genomic Grade Indexet (GGI) [48] pedig 64 ER-pozitív minta alapján fejlesztették ki.

A Recurrence Online (<http://www.recurrenceonline.com/>) egy olyan online rendszer, amellyel automatizált módon eldönthető, hogy melyik terápiás lehetőség – hormon-, kemo- vagy célzott terápia – a legmegfelelőbb a beteg számára. A vizsgálandó tumormintát Affymetrix géncsipekre kell hibridizálni, majd az ennek eredményeképpen létrejövő fájl fel kell tölteni, és a rendszer ezután a számításokat egy központi számítógépen végzi el [49]. A daganat ER-státusza a hormonterápia lehetőségét, a HER2-státusz a célzott terápiás kezelés esélyét, a 21 gén expresszióját használó „RS-calculator” pedig a kemoterápia szükségességét jelzi.

A program egy alternatív osztályozó algoritmust is tartalmaz, amelyben nincsenek előre megadott gének, hanem a géncsipeken lemért összes gén kifejeződését felhasználja. A teljes génexpressziós profil alapján meghatározza, hogy egy korábbi, ~3500 mintát tartalmazó adatbázisban mely betegek mintái hasonlítanak az aktuálisan vizsgált mintához, majd azt számolja ki, hogy mi történt ezekkel a betegekkel [50]. Az

BOADICEA



4. ÁBRA. Rizikó becslése a BOADICEA segítségével

ennek a „Re-training”-nek nevezett módszernek a segítségével adott osztályozás nemcsak hatékonyabb a korábbi, fix géneket tartalmazó osztályozóknál, de hatásos tripla-negatív és HER2-pozitív emlőtumoros betegek túlélésének előrejelzésében is. Mind az „RS-calculator”, mind a „Re-training”

algoritmus az eredményeket grafikus formában és egy le-tölthető jelentésben is elkészíti, amelyek a későbbiekben közvetlenül felhasználhatóak.

RIZIKÓ BECSLÉSE

Összefoglalónk végén meg kell említenünk egy olyan témát is, amely nem közvetlenül az emlőtumoros betegekhez kapcsolódik, hanem a valamilyen rizikótényezővel rendelkező, de még egészséges alanyokat érinti. Már korábban is készültek olyan programok, amelyek a családi háttér adatait egy epidemiológiai adatbázishoz viszonyítva adtak egy becslést az emlőtumor várható kialakulására [51]. Az ehhez a rendszerhez tartozó programot azonban még külön le kell tölteni (<http://www.ems-trials.org/riskevaluator/>), majd a saját számítógépen futtatni.

Egy újabb rendszer a 2785 család adatai alapján készült „Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm” (BOADICEA). A BOADICEA alapja, hogy a rizikót az azt közvetlenül meghatározó legfontosabb BRCA1 és BRCA2 mellett még egy többgénes komponens is befolyásolja. Az erre vonatkozó kombinált modell alapján készített előrejelzés nagy hasonlóságot mutat az epidemiológiai vizsgálatokban kapott megfigyelésekkel [52]. Az interneten keresztül ingyenesen használható, azonban regisztrációhoz kötött BOADICEA program (<http://ccge.medschl.cam.ac.uk/boadicea/boadicea-web-application/>) a számítás első lépéseként bekéri a családtagok klinikai történetét. Ez az általános adatok mellett, mint túlélés, nem stb., az esetleges korábbi

1. TÁBLÁZAT. A cikkben bemutatott online rendszerek áttekintése az első működő verzió megjelenésének évével

ÉV	NÉV	HONLAP	CÉL	REFERENCIA
2001	Adjuvant! Online	https://www.adjuvantonline.com/	Túlélésbecslés	[13]
2007	MSKCC nomogram	http://nomograms.mskcc.org/breast	Túlélésbecslés	[15]
2003	MD Anderson nomogram	http://www.mdanderson.org/education-and-research/resources-for-professionals/clinical-tools-and-resources/clinical-calculators/index.html	Túlélésbecslés	[18]
2010	PREDICT	http://www.predict.nhs.uk/predict.html	Túlélésbecslés	[19]
1995	KEGG	http://www.genome.jp/kegg/	Hálózatelemzés	[39]
2007	DAVID	https://david.ncifcrf.gov/	Hálózatelemzés	[41]
2014	NetGO	http://netgoplot.org/	Hálózatelemzés	[43]
2010	KM-plotter	http://kmplot.com	Biomarker-validáció	[24]
2012	cBioPortal	http://www.cbioportal.org/	Biomarker-validáció	[33]
2015	Genotype to outcome	http://www.g-2-o.com	Túlélésbecslés	[34]
2012	RecurrenceOnline	http://www.recurrenceonline.com/	Teljes diagnosztika	[47]
2004	BOADICEA	http://ccge.medschl.cam.ac.uk/boadicea/boadicea-web-application/	Rizikóbecslés	[52]

tumrok pontos klinikai és patológiai adatait is tartalmazza, mint emlőtumor esetében az ER-, PR- és HER2-státusz. A számítás eredményeképpen a rizikó százalékos értékét, valamint az alaprizikóhoz képest várható emelkedés mértékét mutató grafikont kap vissza a felhasználó, amelyeket külön ki is menthet a saját számítógépére (lásd a 4. ábrán⁴). A BOA-DICEA számításához személyes adatokat is fel kell tölteni a rendszerbe, ezeket azonban a rendszer nem jegyzi meg – a felhasználó kilépése után valamennyi adat törlődik [53].

KITEKINTÉS

A cikkben bemutatott rendszereket az 1. táblázatban foglaltuk össze. Valamennyi online elérhető rendszerben közös, hogy rendkívül könnyű egyrészt a már meglévő adatokat frissíteni, valamint a számítások alapjául szolgáló adatbázisokat kiegészíteni további mintákkal. Ez előrevetíti, hogy a jövőben ezen elemző eszközök hatékonysága a folyamatos frissítéseken keresztül tovább fog javulni.

IRODALOM

- Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012. *CA Cancer J Clin* 65:87–108, 2015
- Sotiriou C, Pusztai L. Gene-expression signatures in breast cancer. *N Engl J Med* 360:790–800, 2009
- Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature* 406:747–752, 2000
- Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 98:10869–10874, 2001
- Cheang MC, Chia SK, Voduc D, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J Natl Cancer Inst* 101:736–750, 2009
- Brenton JD, Carey LA, Ahmed AA, et al. Molecular classification and molecular forecasting of breast cancer: ready for clinical application? *J Clin Oncol* 23:7350–7360, 2005
- Dawood S, Broglio K, Buzdar AU, et al. Prognosis of women with metastatic breast cancer by HER2 status and trastuzumab treatment: an institutional-based review. *J Clin Oncol* 28:92–98, 2010
- Calza S, Hall P, Auer G, et al. Intrinsic molecular signature of breast cancer in a population-based cohort of 412 patients. *Breast Cancer Res* 8:R34, 2006
- Rouzier R, Perou CM, Symmans WF, et al. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. *Clin Cancer Res* 11:5678–5685, 2005
- Herschkwitz JI, Simin K, Weigman VJ, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol* 8:R76, 2007
- Szasz AM, Nemeth Z, Gyorffy B, et al. Identification of a claudin-4 and E-cadherin score to predict prognosis in breast cancer. *Cancer Sci* 102:2248–2254, 2011
- Prat A, Parker JS, Karginova O, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res* 12:R68, 2010
- Ravdin PM, Siminoff LA, Davis GJ, et al. Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. *J Clin Oncol* 19:980–991, 2001
- Olivetto IA, Bajdik CD, Ravdin PM, et al. Population-based validation of the prognostic model ADJUVANT! for early breast cancer. *J Clin Oncol* 23:2716–2725, 2005

Az internetelés elterjedtsége mára Magyarországon is felzárkózott az Európai Unió átlagához (75% a teljes lakosságra vetítve), amely alapján az itt leírt eszközök elérhetősége nálunk is univerzálisnak mondható. A legtöbb eszköz alkalmazásának azonban továbbra is gátja, hogy lényegében valamennyi platform csak angol nyelven érhető el. További nehézség, hogy például az epidemiológiai adatok alapjául szolgáló adatbázisok sem nem magyar, sem nem közép-európai minták felhasználásával készültek, ami különösen az amerikai adatok esetén kérdésessé teheti alkalmazhatóságukat.

Mindezekről függetlenül megállapíthatjuk, hogy a közeljövőben az elmúlt tíz évben megjelent internetalapú diagnosztikai és elemző rendszerek további terjedése várható. Bár az emlő rosszindulatú daganatainak vizsgálata a rendelkezésre álló nagy adatbányák miatt kiemelten sikeres, azonban a közeljövőben más tumortípusok esetén is hasonló paletta megjelenésére számíthatunk.

- Bevilacqua JL, Kattan MW, Fey JV, et al. Doctor, what are my chances of having a positive sentinel node? A validated nomogram for risk estimation. *J Clin Oncol* 25:3670–3679, 2007
- Van Zee KJ, Manasseh DM, Bevilacqua JL, et al. A nomogram for predicting the likelihood of additional nodal metastases in breast cancer patients with a positive sentinel node biopsy. *Ann Surg Oncol* 10:1140–1151, 2003
- Rudloff U, Jacks LM, Goldberg JI, et al. Nomogram for predicting the risk of local recurrence after breast-conserving surgery for ductal carcinoma in situ. *J Clin Oncol* 28:3762–3769, 2010
- Hwang RF, Krishnamurthy S, Hunt KK, et al. Clinicopathologic factors predicting involvement of nonsentinel axillary nodes in women with breast cancer. *Ann Surg Oncol* 10:248–254, 2003
- Wishart GC, Azzato EM, Greenberg DC, et al. PREDICT: a new UK prognostic model that predicts survival following surgery for invasive breast cancer. *Breast Cancer Res* 12:R1, 2010
- Wishart GC, Bajdik CD, Azzato EM, et al. A population-based validation of the prognostic model PREDICT for early breast cancer. *Eur J Surg Oncol* 37:411–417, 2011
- Down SK, Lucas O, Benson JR, et al. Effect of PREDICT on chemotherapy/trastuzumab recommendations in HER2-positive patients with early-stage breast cancer. *Oncol Lett* 8:2757–2761, 2014
- Swanton C, Larkin JM, Gerlinger M, et al. Predictive biomarker discovery through the parallel integration of clinical trial and functional genomics datasets. *Genome Med* 2:53, 2010
- van't Veer LJ, Bernards R. Enabling personalized cancer medicine through analysis of gene-expression patterns. *Nature* 452:564–570, 2008
- Gyorffy B, Lanczky A, Eklund AC, et al. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res Treat* 123:725–731, 2010
- Szasz AM, Lanczky A, Nagy A, et al. 2016. Cross-validation of survival associated biomarkers in gastric cancer using transcriptomic data of 1,065 patients. *Oncotarget* 2016, DOI: 10.18632/oncotarget.10337
- Gyorffy B, Surowiak P, Budczies J, Lanczky A. Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. *PLoS One* 8:e82241, 2013
- Li Q, Birkbak NJ, Gyorffy B, et al. Jetset: selecting the optimal microarray probe set to represent a gene. *BMC Bioinformatics* 12:474, 2011
- Gyorffy B, Schafer R. Meta-analysis of gene expression profiles related to relapse-free survival in 1,079 breast cancer patients. *Breast Cancer Res Treat* 118:433–441, 2009
- Heber S, Sick B. Quality assessment of Affymetrix GeneChip data. *OMICS* 10:358–368, 2006

⁴Hozzáférés időpontja: 2015. IX. 26.

30. Mihaly Z, Kormos M, Lanczky A, et al. A meta-analysis of gene expression-based biomarkers predicting outcome after tamoxifen treatment in breast cancer. *Breast Cancer Res Treat* 140:219–232, 2013
31. Gyorffy B, Bottai G, Lehmann-Che J, et al. TP53 mutation-correlated genes predict the risk of tumor relapse and identify MPS1 as a potential therapeutic kinase in TP53-mutated breast cancers. *Mol Oncol* 8:508–519, 2014
32. Lánzky A, Nagy Á, Bottai G, et al. miRpower: a web-tool to validate hip-erasszociativitás miRNAs utilizing expression data from 2,178 breast cancer patients *Breast Cancer Res Treat*, 160:439–446, 2016
33. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2:401–404, 2012
34. Pongor LS, Kormos M, Hatzis C, et al. A genome-wide approach to link genotype to clinical outcome by utilizing next generation sequencing and gene chip data of 6,697 breast cancer patients. *Genome Med* 7:104, 2015
35. Nagy Á, Pongor LS, Szabó A, et al. KRAS driven expression signature has prognostic power superior to mutation status in non-small cell lung cancer. *Int J Cancer*, 2016, in press.
36. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature* 458:719–724, 2009
37. International HapMap Consortium. The International HapMap Project. *Nature* 426:789–796, 2003
38. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 6:95–108, 2005
39. Kanehisa M, Goto S, Furumichi M, et al. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res* 38(Database issue):D355–360, 2010
40. Kanehisa M, Goto S, Sato Y, et al. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 40(Database issue):D109–114, 2012
41. Huang DW, Sherman BT, Tan Q, et al. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res* 35(Web Server issue):W169–175, 2007
42. Huang DW, Sherman BT, Tan Q, et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol* 8:R183, 2007
43. Shi W, Balazs B, Gyorffy B, et al. Combined analysis of gene expression, DNA copy number, and mutation profiling data to display biological process anomalies in individual breast cancers. *Breast Cancer Res Treat* 144:561–568, 2014
44. Penzvalto Z, Mihaly Z, Gyorffy B. Génexpresszió mérésén alapuló multigénes prognosztikai és prediktív előrejelzés emlőtumorkban. *Magy Onkol* 53:351–359, 2009
45. Gyorffy B, Hatzis C, Sanft T, et al. Multigene prognostic tests in breast cancer: past, present, future. *Breast Cancer Res* 17:11, 2015
46. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 347:1999–2009, 2002
47. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 351:2817–2826, 2004
48. Sotiropoulos C, Wirapati P, Loi S, et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst* 98:262–272, 2006
49. Gyorffy B, Benke Z, Lanczky A, et al. RecurrenceOnline: an online analysis tool to determine breast cancer recurrence and hormone receptor status using microarray data. *Breast Cancer Res Treat* 132:1025–1034, 2012
50. Gyorffy B, Karn T, Sztupinszki Z, et al. Dynamic classification using case-specific training cohorts outperforms static gene expression signatures in breast cancer. *Int J Cancer* 136:2091–2098, 2015
51. Anderson H, Bladstrom A, Olsson H, et al. Familial breast and ovarian cancer: a Swedish population-based register study. *Am J Epidemiol* 152:1154–1163, 2000
52. Antoniou AC, Pharoah PP, Smith P, et al. The BOADICEA model of genetic susceptibility to breast and ovarian cancer. *Br J Cancer* 91:1580–1590, 2004
53. Cunningham AP, Antoniou AC, Easton DF. Clinical software development for the Web: lessons learned from the BOADICEA project. *BMC Med Inform Decis Mak* 12:30, 2012

FELHÍVÁS

Kedves Kollegák, ezúton szeretnénk felhívni figyelmüket, hogy a

XXXII. MOT Kongresszus

(A Magyar Onkológusok Társasága®
XXXII. Kongresszusa)

2017. november 16–18.

között, a debreceni Kölcsey Központban kerül megrendezésre

