

# Hasznos tesztek az új prognosztikai markerek klinikai értékének felmérésére: a szívelégtelenség példája

Jenei Zsigmond Máté dr. ■ Prohászka Zoltán dr.

Semmelweis Egyetem, Általános Orvostudományi Kar, III. Belgyógyászati Klinika, Kutatólaboratórium, Budapest

*Bevezetés:* Napjaink klinikai kutatásainak egyik fő iránya a megbetegedések hátterében álló kockázati tényezők azonosítása. Közlemények százai számolnak be „szignifikáns” és „független” prognosztikai faktorokról különböző humán megbetegedésekben, azonban ezek egy részében nem vagy nem megfelelően vizsgálták, hogy az új prognosztikai faktor javította-e, és ha igen, milyen mértékben az addig ismert prognosztikai modellt. A legújabb statisztikai módszertani ajánlások szerint az úgynevezett reklassifikációs analízissel a fenti kérdést érdemben vizsgálni lehet. *Célkitűzés:* A reklassifikációs analízis kivitelezésére több módszer is van, a közleményben a szerzők ezek alkalmazását saját, korábban publikált vizsgálataik újraelemzésével mutatják be. *Módszer:* Két marker, a vörösvértestátmérő-eloszlás szélessége és a szérum-hősokkfehérje-70 prognosztikai szerepét vizsgálták krónikus szívelégtelenségben szenvedő betegek körében. Korábban publikált eredményeik szerint mind a vörösvértestátmérő-eloszlás szélessége, mind a hősokkfehérje-70 szignifikáns, független prognosztikai markernek bizonyult többváltozós Cox-regressziós vizsgálatok alapján. Mindkét esetben újraértékelték a markerek szerepét reklassifikációs tesztekkel. *Eredmények:* A szívelégtelen betegek prognosztikai modelljének diszkriminatív képessége lényegesen javult a korábbi modellhez képest, ha a vörösvértestátmérő-eloszlás szélességével egészítették ki a modellt, míg a hősokkfehérje-70 esetén ez nem volt egyértelmű. *Következtetések:* Az új prognosztikai faktorokat kritikusán kell kezelni mindaddig, míg alkalmas vizsgálatokkal elemzésre és bizonyításra nem kerül, mekkora a valós klinikai haszon, amely a marker már ismert prognosztikai modellhez való hozzáadásából származik. A hasznosságot a prognosztikai modell javulása során tesztelhetjük a reklassifikáció módszereivel. Orv. Hetil., 2013, 154, 1374–1380.

**Kulcsszavak:** prognosztikai marker, kockázatbecslés, kimenetel, reklassifikáció, szívelégtelenség

## Useful tests to assess the clinical usefulness of new prognostic markers: the example of heart failure

*Introduction:* Identification of risk factors is one of the most frequent questions in medical research currently. Several reports showed “significant” and “independent” prognostic factors in a variety of human conditions, however, those were not tested about predictive information in addition to standard risk markers. Recently novel statistical approaches (reclassification) have been developed to test the performance and usefulness of new risk factors and prognostic markers. There are several established methods to test the prognostic models. *Aim:* The aim of this work was to present the application of these novel statistical approaches by re-analyzing previously reported results of the authors. *Method:* The authors analyzed the prognostic role of two markers: red cell distribution width and heat shock protein 70 in patients with heart failure. Using Cox regression analyses the authors have reported previously that both markers are independent predictors. In the present study they re-analyzed the role of red cell distribution width and heat shock protein 70 by reclassification tests. *Results:* Incorporating red cell distribution width to the reference model the authors found a significant improvement in discrimination. However, the reclassification analysis provided ambiguous results with heat shock protein 70. *Conclusions:* Interpretation of results on new prognostic factors has to be done carefully, and appropriate reclassification approaches may help to confirm clinical usefulness only. Orv. Hetil., 2013, 154, 1374–1380.

**Keywords:** prognostic marker, risk prediction, outcome, reclassification, heart failure

(Beérkezett: 2013. június 27.; elfogadva: 2013. július 11.)

### Rövidítések

AUC = (area under curve) görbe alatti terület; HSP70 = (heat shock protein 70) hősokkfehérje-70; IDI = integrated discrimination improvement; NRI = net reclassification improvement; RDW = (red cell distribution width) vörösvértestátmérő-eloszlás szélessége; ROC = receiver operating characteristic

Ötven évvel ezelőtt, a Framingham Study eredményeinek publikálásakor kapott az orvosi irodalomban a rizikófaktor fogalma nagyobb figyelmet [1], és ezt követően dolgozták ki a cardiovascularis betegségek kifejlődésének előrejelzésére az első többváltozós prognosztikai modellt [2]. A prognosztikai modell olyan eszköz, amelynek a klinikai döntéshozatalban van szerepe. A prognosztikai modell alapján egyszerűen használható kalkulátorokat lehet létrehozni, amelyek segítségével egyes eseményeket (például betegség kialakulása, halálozás) kellő pontossággal előre lehet jelezni. Ilyen például az intenzív betegellátásban használt APACHE II score, amely a túlélésre ad információt a klinikai adatok alapján, vagy a CHA2DS2-VASc score, amely pitvarfibrilláló betegek stroke-rizikójára ad valószínűségi becslést.

Az ezzel kapcsolatos kutatások azóta is folyamatosan foglalkoznak az újabb rizikófaktorok meghatározásával és a faktorok megfelelő rizikóbecslő modellekbe történő beépítésével. Míg gyakran az elsők között azonosított („erős”) rizikótényezők klinikai hasznossága a döntéshozatalban megkérdőjelezhetetlen (például magas székumkoleszterin-szint és cardiovascularis halálozás) [3], addig az új és még újabb rizikófaktorok vagy rizikómarkerek (például CRP és cardiovascularis halálozás) [4, 5] klinikai értéke nem azonnal nyilvánvaló. Közlemények százai számolnak be újabb és újabb szignifikáns, független prognosztikai és prediktív faktorokról különböző humán megbetegedések kapcsán, azonban nem vizsgálták vagy nem megfelelően vizsgálták, hogy az új prognosztikai vagy rizikófaktorok javítják-e, és ha igen, milyen mértékben az addig ismert modellt, vagyis pontosabb becslést adnak-e a prognózisra, és ezáltal érdemben képesek-e javítani a klinikai döntéshozatalt. Valójában a rengeteg potenciálisan jó, az adott betegségben zajló patogenetikai eseményekre rámutató marker közül nagyon kevés az, amelynek igazolható a valódi klinikai haszna [6]. Minél több tényezőt vizsgálunk, annál több szignifikáns prognosztikai markert találhatunk, és ez különösen a „nagy áteresztőképességű” genomikai és proteomikai kutatások esetén igaz [7].

A gyógyítás szempontjából az események (kimenetel) bekövetkeztenek megjósolásának a legnagyobb a jelentősége, és ebben az értelemben az esemény lehet diagnózis vagy túlélési adat. Rizikófaktor az a tényező, amely oklag, követhető patofiziológiai láncolaton keresztül hozzájárul egy betegség kifejlődéséhez (például dohányzás és tüdőrák, koraszülöttség és krónikus tüdőbetegség). A már kialakult betegség további kimeneteléről adnak információt a prognosztikai és a prediktív markerek.

Definíció szerint prognosztikai marker az, amely a beteg túlélésével, a betegség kimenetelével mutat kapcsolatot, függetlenül az alkalmazott terápiától (például általános állapot, metasztázisok, tumorstádium stb.), a prediktív marker ezzel szemben a tervezett terápia hatékonyságáról ad információt (például ösztrogén- és progeszteron-receptor-expresszió mértéke emlőrákban). Ezen definíciók azonban még a mai nemzetközi irodalomban is gyakran helytelenül, felcserélve kerülnek alkalmazásra. Az utóbbi években többféle statisztikai eljárást dolgoztak ki és alkalmaztak a rizikófaktorok azonosítására és hatásuk tesztelésére, ezek közül napjainkban az orvosi vizsgálatokban a többváltozós Cox-regressziós eljárás az egyik legelterjedtebb módszer [8]. A Cox-regressziós elemzés a túlélési vizsgálatok csoportjába tartozik, és teszteli, hogy a vizsgált tényező összefüggést mutat-e a túléléssel. Többváltozós (adjusztált) modelltől abban az esetben beszélünk, amikor ismert, túlélést befolyásoló tényezőkkel együtt (például életkor, nem) elemezzük a vizsgált faktor hatását a túlélésre. Mára azonban körvonalazódott, hogy ez a módszer valójában korlátozott értékű az új markerek valódi klinikai hasznosságával kapcsolatban [9]. Továbbá az elterjedten alkalmazott c-statisztika módszeréről – amely a ROC-görbe alatti terület növekedését (azaz a modell diszkriminatív képességét) teszteli – kiderült, hogy a klinikailag fontos rizikófaktorok hatását tévesen alulbecsli [10, 11]. Mivel a Cox-regresszió és a c-statisztika nem megfelelően alkalmas módszerek, az új potenciális biomarkerek klinikai hasznosságának tesztelésére az elmúlt években új, reklasszifikáción alapuló statisztikai módszereket vezettek be, amelyek fokozatos elterjedésének lehetünk szemtanúi [12, 13].

Egy prognosztikai modellel szemben támasztott alapkövetelmény az, hogy megbízhatóan elkülönítse, diszkriminálja azokat a betegeket, akikben az esemény bekövetkezik, azoktól, akikben nem. Ha egy új prognosztikai modell a korábbihoz képest precízebben diszkriminál, pontosabb predikciót ad, akkor kisebb a hiba, vagyis az eltérés az előre jelzett vagy kalkulált rizikó (predicted risk) és a valóság (observed risk) között. Az eltérés jellemzésére használhatjuk a fals pozitív rátát (I. fajú hiba, azaz  $1 - \text{mínusz specifititás}$ ) és a fals negatív rátát (II. fajú hiba, azaz  $1 - \text{mínusz szenzitivitás}$ ). Azt a folyamatot, amikor a modell besorolja a beteget egy előre meghatározott rizikó kategóriába a kalkulált rizikó (predicted risk) alapján, klasszifikációnak nevezzük. Ha a besorolást az új prognosztikai modell segítségével újra elvégezzük, reklasszifikációt végzünk. A reklasszifikáción alapuló módszerek azt mutatják meg, hogy az új modell a régihez képest a betegek mekkora hányadát sorolja be jobban a meghatározott kategóriákba, vagyis mennyivel több a jól klasszifikált beteg. Az új reklasszifikációs módszerek, nevezetesen az NRI (net reclassification improvement) és az IDI (integrated discrimination improvement) segítségével a korábbi módszerekhez képest jobban megítélhető, hogy az új tényező (rizikófaktor

vagy prognosztikai marker) modellbe építése hogyan befolyásolja a prognosztikai modell diszkriminatív erejét, vagyis mennyire hasznos az új tényező, és az új modell mennyivel jobb, mint a korábbi. Az új rizikófaktorok kritikus értékeléséhez *Hlatky és munkatársai* készítettek nemrég ajánlást [12]. Jelen munkánk célja az volt, hogy ezen elemzéseket saját vizsgálataink újraszámolásának példáján bemutassuk és rávilágítsunk a lehetséges buktatókra. Az elemzések során munkacsoportunk korábban, a szívelégtelenséggel kapcsolatban publikált eredményeit értékeltük újra a fenti statisztikai módszerek segítségével.

## Módszer

Két korábbi cikkünkben két prognosztikai markert vizsgáltunk krónikus szívelégtelenségben szenvedő betegek körében, nevezetesen a vörösvértestátmérő-eloszlás szélességét (RDW) és a szérum-hőssokkfehérje-70-et (HSP70). A tanulmányainkban a többváltozós Cox regressziós elemzések azt mutatták, hogy mind az RDW, mind a HSP70 szignifikáns, független prognosztikai marker krónikus szívelégtelenségben szenvedő betegek túlélésére vonatkozóan. Az első elemzésben az RDW kapcsolatot mutatott a betegek egyéves túlélésével [14], míg a másodikban a HSP70 mutatott kapcsolatot az öt éves túléléssel [15].

Az elemzést 195 krónikus szívelégtelen betegen végeztük, akik 2005. február és 2007. április között kerültek bevonásra a Semmelweis Egyetem III. Belgyógyászati Klinikáján konzekutív módon. A beválasztási kritériumok között szerepelt a manifeszt bal kamrai szisztolés szívelégtelenség tüneteinek megléte és a 45%-nál kisebb bal kamrai ejekciós frakció. Kizárási kritériumok a malignus alapbetegség és az infekció voltak. A vizsgálati protokoll szerint a bevonás során minden beteg fizikális és transthoracalis echokardiográfiás vizsgálaton esett át. A vérvételek reggel, 12 óras éhezés után, az antecubitalis véna punkciójával történtek. A betegek követési ideje 12, illetve 60 hónap volt, a követés alatt rögzítettük az összmortalitást.

A vizsgálatok részletes leírása, az alkalmazott mérési és analitikai módszerek ismertetése a korábbi közleményekben található [14, 15].

## Statisztikai elemzés

Statisztikai számításokhoz az R nyílt forráskódú programot (R version 2.15.0, 2012-03-30) használtuk a PredictABEL [16] és pROC [17] statisztikai programcsomagok alkalmazásával, kivéve a Cox-regressziót, amelyet SPSS v13.0 (SPSS Inc., Chicago, IL, Amerikai Egyesült Államok) programmal végeztük. Az RDW-t vizsgáló közlemény Cox-regressziós modelljében alkalmazott adjusztált változók, amelyeket együttesen tekintettük a referenciamodellnek, a következők voltak: életkor, glomerularis filtrációs ráta, hemoglobinn, testtömegindex, diasztolés vérnyomás, szérumnátriumszint és NT-proBNP.

A HSP70-et vizsgáló cikk Cox-regressziós modelljében alkalmazott adjusztált változók, amelyeket szintén együttesen tekintettük a referenciamodellnek, a következők voltak: életkor, nem, testtömegindex, szérumkreatinin és NT-proBNP. A túlélési modell általános jellemzésére Nagelkerke-féle  $R^2$ -et és Brier-score-t használtunk, míg a diszkriminációs képességet a ROC-görbe alatti terület és a diszkriminációs meredekség (slope) alapján számoltuk, valamint ábrázoltuk a diszkriminációs képességet ROC-görbén és Box-ploton is a legújabb ajánlásnak megfelelően [13]. A kalibrációt Hosmer–Lemeshow-teszt segítségével elemeztük, és grafikusán is bemutattuk. Az alapmodellhez tartozó és a HSP70-nel kibővített modellhez tartozó ROC-görbéket DeLong-teszttel hasonlítottuk össze [18]. A reklassifikációs elemzések közül a vágópont nélküli NRI-t (continuous NRI) és IDI-t számoltuk ki, és megadtuk a 95%-os konfidenciaintervallumot, valamint a hozzá tartozó p-értéket. Minden esetben az eredeti közleményben szereplő követési időszak adatait vettük figyelembe a reklassifikációs elemzés során, és a  $p < 0,05$ -ot tekintettük statisztikailag szignifikánsnak.

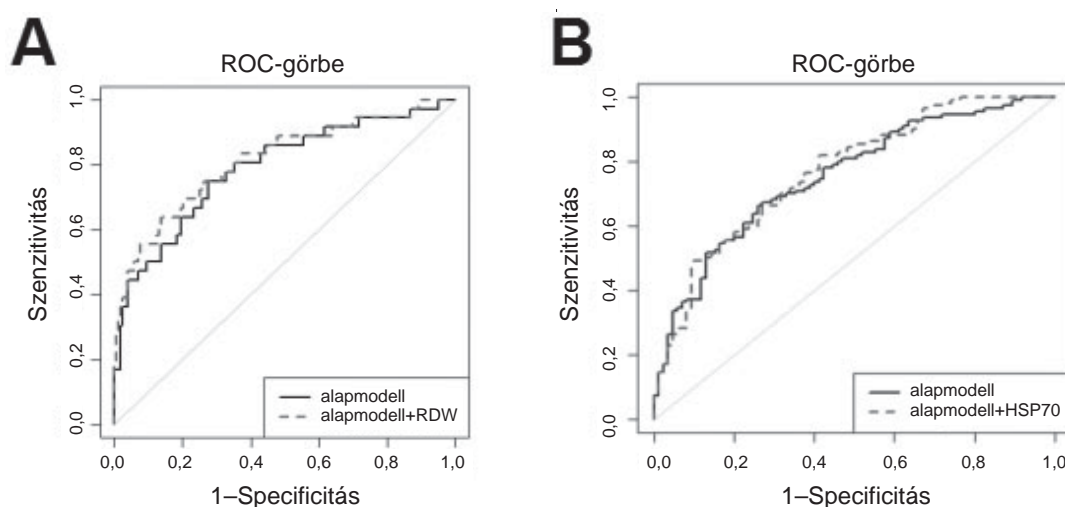
## Eredmények

Mindkét vizsgálatot újraértékeltek a reklassifikáció módszereivel, és elvégeztük a modellek részletes jellemzését, amelynek eredménye az *1. táblázatban* látható. A Cox-regresszió szignifikáns, független prediktornak jelezte mind az RDW-t, mind a HSP70-et ( $p < 0,0001$  és  $p = 0,032$ , *1. táblázat*). Azonban egyik esetben sem nőtt szignifikáns mértékben a ROC-görbe alatti terület (0,790-ról 0,812-re, illetve 0,754-ről 0,769-re változott), amikor az RDW-vel, illetve a HSP70-nel egészítettük ki a referenciamodellt (a DeLong-teszt nem volt szignifikáns, *1. táblázat*). Az *1. ábrán* a két elemzéshez tartozó ROC-görbék kerültek bemutatásra. Mindkét esetben növekedett a Nagelkerke-féle  $R^2$ , és ezzel párhuzamosan csökkent a Brier-score, valamint az NRI-számolás során szignifikáns javulást észleltünk, amikor a HSP70-nel kiegészítettük a referenciamodellt (*1. táblázat*). Habár az IDI értéke is mindkét vizsgálatban pozitív volt, csak az RDW-t vizsgáló elemzésben találtunk statisztikailag igazolható növekedést ( $p < 0,0495$ , *1. táblázat*). Az IDI értékét a „discrimination slope”-ok különbsége adja. Ez a *2. ábrán* vizuálisan is megítélhető. Látható a *2. ábra A és B paneljén*, hogy míg a referenciamodell RDW-vel való kibővítése után a túlélőket és nem túlélőket jobban elkülöníti egymástól (a két átlagérték közti különbség megnő), azaz jobban diszkriminál, addig a HSP70 esetén a grafikus ábrázolás közel azonos mértékű diszkriminációt mutat (*2. ábra C és D panel*). A kalibráció vizsgálatok azt tapasztaltuk, hogy az RDW-t vizsgáló elemzésben a kibővített modell jónak bizonyult, mivel a Hosmer–Lemeshow-teszt nem szignifikáns (*1. táblázat*), és a grafikus ábrázolás is a 45 fokos kalibrációs egyeneshez közeli elrendeződést mutat (*3. ábra A és B panel*). Ezzel szemben a HSP70-et vizs-

**1. táblázat** | A referenciamodell (alapmodell) és a kibővített modell (alapmodell+RDW, alapmodell+HSP70) fő statisztikai jellemzői láthatóak a táblázatban. A szignifikáns ( $p < 0,05$ ) értékeket félkövér betűtípussal jelöltük. A statisztikai tesztekéről a bevezetőben talál részletes információt az olvasó

	RDW		HSP70	
	Alapmodell	Alapmodell+RDW	Alapmodell	Alapmodell+HSP70
Cox-regresszió $\chi^2$	59,328	73,937	74,863	78,447
Cox-regresszió p		<b>&lt;0,0001</b>		<b>0,0320</b>
Hosmer–Lemeshow-teszt $\chi^2$	1,635	0,921	9,710	19,400
p (Hosmer–Lemeshow-teszt)	0,9902	0,9987	0,2860	<b>0,0129</b>
Nagelkerke-féle $R^2$	0,3020	0,3625	0,2514	0,2693
Brier-score	0,1144	0,1050	0,1988	0,1944
C-statisztika (AUC)	0,790	0,812	0,754	0,769
95% CI (AUC)	0,702–0,878	0,726–0,898	0,687–0,822	0,703–0,834
p (DeLong-teszt)		0,2235		0,2973
NRI (continuous)		0,4539		0,3144
95% CI (NRI)		0,0921–0,8156		0,0314–0,5975
p (NRI)		<b>0,01393</b>		<b>0,0295</b>
IDI		0,057		0,0156
95% CI (IDI)		0,0001–0,1139		–0,0016–0,0328
p (IDI)		<b>0,0495</b>		<b>0,0746</b>

AUC = (area under curve) görbe alatti terület; 95% CI = (95% confidence interval) 95%-os megbízhatósági tartomány; NRI = net reclassification improvement; IDI = integrated discrimination improvement; RDW = (red cell distribution width) vörösvértestátmérő-eloszlás szélessége; HSP70 = (heat-shock protein 70) hőszokkfehérje-70.



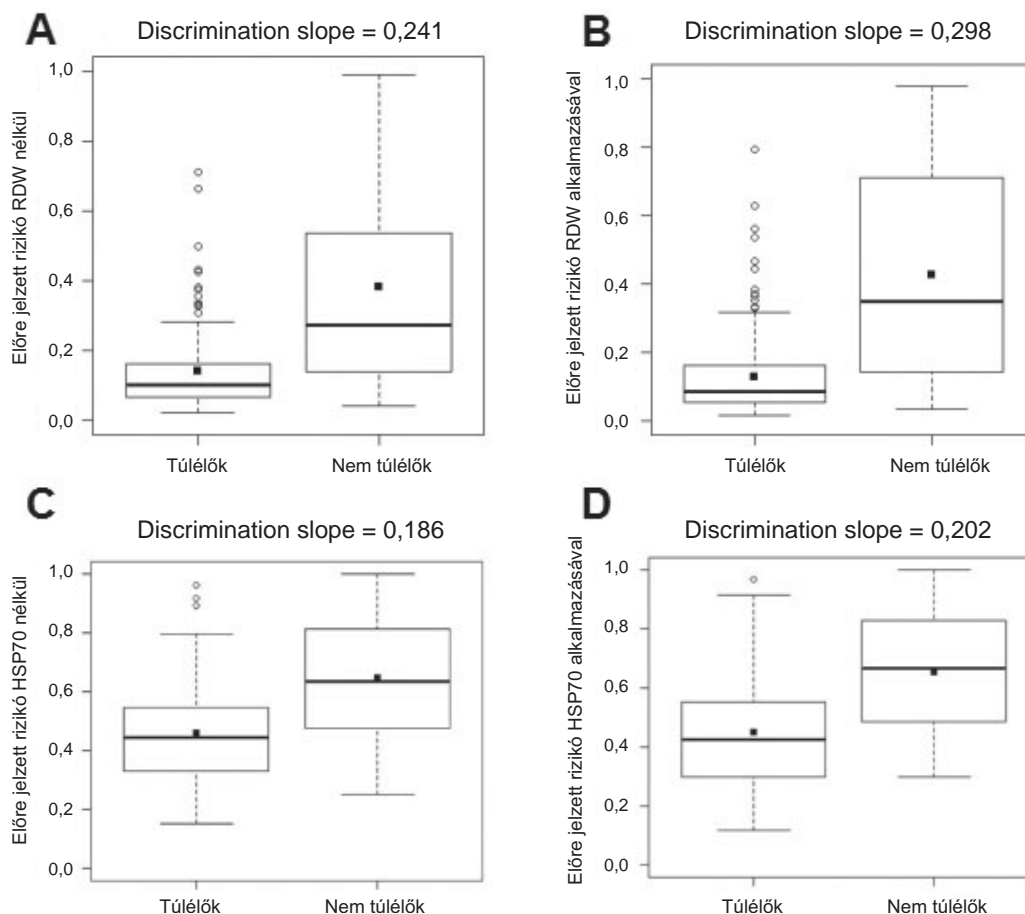
**1. ábra** | Az alap prognosztikai modellekhez (alapmodell) és a kibővített modellekhez (alapmodell+RDW, alapmodell+HSP70) tartozó ROC (receiver operating characteristic) görbék. Az A panelen az RDW-, a B panelen a HSP70-analízishez tartozó görbék kerültek bemutatásra

gáló elemzésben a Hosmer–Lemeshow-teszt már szignifikáns ( $p=0,0129$ ), és a kalibrációs egyenestől jelentős eltéréseket látunk a grafikus ábrázolás során (3. ábra C és D panel). Az eredményeket összesítve láthatjuk, hogy míg az RDW alkalmazásának egyértelműen jelentős a klinikai haszna szívelégtelen betegek hosszú távú prognózisának becslésében, addig a HSP70, mint szignifikáns rizikómarker, klinikai jelentősége kérdéses.

## Megbeszélés

Vizsgálatunk eredménye felhívja a figyelmet arra, hogy a klinikai kutatások során feltárt új prognosztikai faktorokat gondosan meg kell vizsgálni és kritikusán kell kezelni mindaddig, míg bizonyításra nem kerül, mekkora a valódi klinikai haszon, amelyet az adott marker jelenthet. Ez a hasznosság az IDI- és NRI-módszerekkel vizsgálható, az eredmény számszerűsíthető, összehasonlítható.





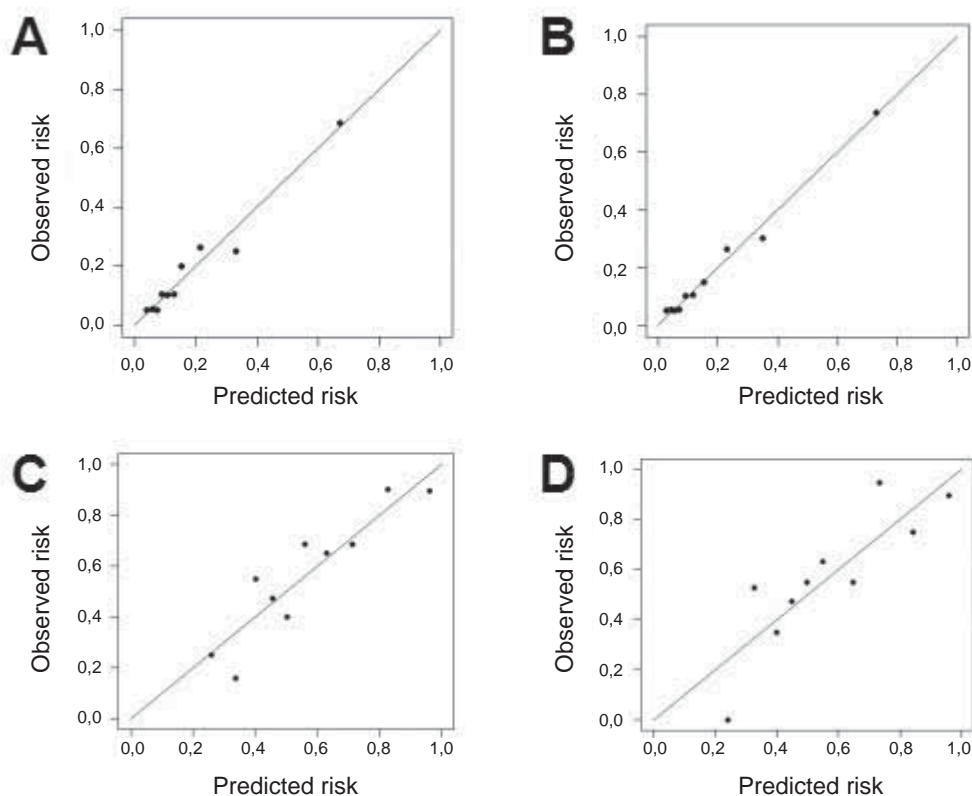
2. ábra

Az alap prognosztikai modellekhez (A, C) és a kibővített prognosztikai modellekhez (B, D) tartozó előre jelzett kockázat (predicted risk) Box-plot ábrázolása. Az A és B panelen az RDW-elemzéshez, a C és D panelen a HSP70-elemzéshez tartozó diszkriminációs jellemzők kerültek bemutatásra, Box = 25–75 percentilis (felső–alsó kvartilis, azaz interkvartilis-tartomány – IQR = interquartile range), pont a Box-ban = átlag, vonal a Box-ban = medián, kapocs = minimum–maximum érték, karika = kilógó érték (1,5 IQR-nél nagyobb eltérés a kvartilisektől). Az ábra fölött az elemzéshez tartozó „discrimination slope”-értékeket tüntettük fel. A „discrimination slope” az előre jelzett rizikó átlagértékének különbsége abban a két csoportban, ahol az esemény (halálozás) bekövetkezett (nem túlélő), illetve ahol nem következett be (túlélő). A „discrimination slope”-ok különbsége adja az IDI értékét

Az IDI- és NRI-elemzés nagyon gyorsan elterjedt az orvosi irodalomban, és nélkülözhetetlen eszközzé kezdett válni napjaink prognosztikai modelleket feltáró vizsgálatainak [6, 13].

Az NRI- vagy „categorical NRI”-elemzés lényege egy komplex kontingenciátábla vagy más néven reklassifikációs tábla, amelyben a korábbi modell alapján (előzetesen meghatározott vágópont [cutoff] szerint) rizikó kategóriákba besorolt betegeket az új modell szerint újra osztályozzuk, és a két modell közti eltéréseket teszteljük. Az NRI módszere a reklassifikációs tábla alapján szám-szerűsíti, hogy a betegek mekkora hányada lesz az új módszer szerint jobban besorolva [8]. Az NRI  $-1$  és  $+1$  között vehet fel értékeket. Amennyiben az új modell jobban diszkriminál, az NRI pozitív értéket vesz fel, fordított esetben negatívát. Ha nincs különbség,  $NRI=0$ . A reklassifikációs táblán alapuló számolásoknak azonban megvan az a hátránya, hogy az eredmény jelentősen függ az előzetesen kiválasztott kategóriáktól (a vágópont-értékektől) [19]. Ezért a fenti gondolatmenetet és elemzést továbbfejlesztették, és létrehozták az NRI két általá-

nosított, vágópontok alkalmazásától független formáját, az IDI-t és a folytonos (continuous) NRI-t. Az IDI-elemzés során előzetesen meghatározott vágópont nélkül számolunk, vagyis az eljárás nem igényel önkényes kategóriabesorolást. Az IDI értéke becslhető Yates szerint a referencia és az új modellben kalkulált „discrimination slope”-ok különbsége alapján [20]. Az ehhez szükséges „discrimination slope” számolása egyszerű matematikai művelet. Az előre jelzett rizikó átlagértékét kell meghatározni abban a két csoportban, ahol az esemény bekövetkezett, illetve ahol nem, és ezen átlagok különbsége adja az adott modellre jellemző „discrimination slope”-ot. Az IDI szintén  $-1$  és  $+1$  között vehet fel értékeket. Amennyiben az új modell jobban diszkriminál, az IDI pozitív értéket vesz fel, fordított esetben negatívát. Ha nincs különbség,  $IDI=0$ . Ehhez a módszerhez nagyon hasonló a vágópont nélküli NRI is, amelyet az irodalomban „continuous NRI” néven találunk, és ez utóbbi alkalmazása javasolt, szemben a „categorical NRI”-vel, mert kiküszöböli a vágópont okozta torzítást [21].



3. ábra

Az A és B panelen az RDW-elemzés, a C és D panelen a HSP70-elemzéshez tartozó kalibrációs egyenesek. Az alap prognosztikai modellek kalibrációját az A és C panelen, míg a kibővített prognosztikai modellek kalibrációját a B és D panelen ábrázoltuk. Az x tengelyen az előre jelzett vagy kalkulált kockázat (predicted risk), míg az y tengelyen a megfigyelt kockázat (observed risk) szerepel

Míg a modell általános jellemzésére a Nagelkerke-féle  $R^2$  és a Brier-score használata általánosan elfogadott, addig a kalibrációra a Hosmer–Lemeshow-teszt, a diszkriminációra pedig a c-statisztika alkalmas [13]. A kalibráció mutatja meg, hogy a predikció és a valódi kimenetel hogyan függ össze [22]. Ha a Hosmer–Lemeshow-teszt szignifikáns, az azt jelzi, hogy a modell nem alkalmas prognózisbecslésre. Grafikusán ábrázolva a kalibrációt, az origón átmenő 45 fokos egyenesre esnek a pontok tökéletes predikció esetén [13]. Definíció szerint a Brier-score 0 ideális modellben és 0,25 a teljesen „noninformatív” modellben. Ezzel szemben a Nagelkerke-féle  $R^2$  0 és 1 között vesz fel értékeket, ahol az 1 jelenti a legjobb modellt.

Ezek alapján megállapíthatjuk, hogy mindkét vizsgálatban javult a modell, amikor új markerrel (RDW, HSP70) egészítettük ki a referenciamodellt, hiszen növekedett a Nagelkerke-féle  $R^2$ , és ezzel párhuzamosan csökkent a Brier-score (1. táblázat). A bemutatott elemzés megerősítette azt, hogy az RDW valóban fontos a szívelégtelen betegek prognózisbecslésében. Az NRI és az IDI szignifikáns értéke alapján állíthatjuk, hogy az RDW mérésével a szívelégtelen betegek túlélését jobban előre lehet jelezni. A HSP70 példáján bemutattuk azt is, hogy hiába szignifikáns egy adott prognosztikai marker az adjusztált Cox-regresszió-elemzésben, az még nem jelenti azt, hogy ebből valós klinikai előny, pontosabb

rizikóbecslés is származik. Az észlelt NRI-növekedés azt jelzi, hogy a HSP70-nel kibővített modell jobban diszkriminál. Azonban ezt cáfolja az IDI-elemzés, amelyben a növekedés nem volt szignifikáns. További független vizsgálatok szükségesek, amelyek alapján egyértelműen el lehet majd dönteni a hősokkfehérje klinikai jelentőségét a szívelégtelenség prognózisában. Eredményeink megerősítik továbbá azt a megfigyelést [10, 11], hogy a görbe alatti terület (AUC) növekedése nem kellően érzékeny módszer, hiszen a DeLong-teszt, amely a AUC-növekedést teszteli, egyik elemzésben sem volt szignifikáns, míg az NRI mindkét esetben szignifikáns javulást jelzett a modell diszkriminatív képességét tekintve.

## Következtetések

Az új prognosztikai faktorokat kritikusan kell kezelni mindaddig, amíg bizonyításra nem kerül, mekkora a valódi klinikai haszon, amely az adott marker vizsgálatából származhat. A hasznosságot a prognosztikai modellen történő javítás során tesztelhetjük a reklassifikáció módszereivel.

## Köszönetnyilvánítás

Köszönetet mondunk a III. Belgyógyászati Klinika (Dr. Borgulya Gábor, Dr. Förhécz Zsolt, Dr. Gombos Tímea PhD, Dr. Jánoskúti Lívía

PhD, Dr. Karádi István DSc, Dr. Pozsonyi Zoltán) munkatársainak, akik hozzájárulása nélkül nem készülhettek volna el a jelen elemzés alapját képező tanulmányok.

## Irodalom

- [1] Kannel, W. B., Dawber, T. R., Kagan, A., et al.: Factors of risk in the development of coronary heart disease – six year follow-up experience. The Framingham Study. *Ann. Intern. Med.*, 1961, 55, 33–50.
- [2] Kannel, W. B., McGee, D., Gordon, T.: A general cardiovascular risk profile: the Framingham Study. *Am. J. Cardiol.*, 1976, 38, 46–51.
- [3] Martin, M. J., Hulley, S. B., Browner, W. S., et al.: Serum cholesterol, blood pressure, and mortality: implications from a cohort of 361,662 men. *Lancet*, 1986, 2 (8513), 933–936.
- [4] Abd, T. T., Eapen, D. J., Bajpai, A., et al.: The role of C-reactive protein as a risk predictor of coronary atherosclerosis: implications from the JUPITER trial. *Curr. Atheroscler. Rep.*, 2011, 13, 154–161.
- [5] Rietzschel, E., De Buyzere, M.: High-sensitive C-reactive protein: universal prognostic and causative biomarker in heart disease? *Biomark. Med.*, 2012, 6, 19–34.
- [6] Sundström, J., Byberg, L., Gedeberg, R., et al.: Useful tests of usefulness of new risk factors: tools for assessing reclassification and discrimination. *Scand. J. Public Health*, 2011, 39, 439–441.
- [7] Schumacher, M., Binder, H., Gerd, T.: Assessment of survival prediction models based on microarray data. *Bioinformatics*, 2007, 23, 1768–1774.
- [8] Pencina, M. J., D'Agostino, R. B. Sr., D'Agostino, R. B. Jr., et al.: Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat. Med.*, 2008, 27, 157–172.
- [9] Uno, H., Tian, L., Cai, T., et al.: A unified inference procedure for a class of measures to assess improvement in risk prediction systems with survival data. *Stat. Med.*, 2013, 32, 2430–2442.
- [10] Pepe, M. S., Janes, H., Longton, G., et al.: Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am. J. Epidemiol.*, 2004, 159, 882–890.
- [11] Zethelius, B., Berglund, L., Sundström, J., et al.: Use of multiple biomarkers to improve the prediction of death from cardiovascular causes. *N. Engl. J. Med.*, 2008, 358, 2107–2116.
- [12] Hlatky, M. A., Greenland, P., Arnett, D. K., et al.: Criteria for evaluation of novel markers of cardiovascular risk: a scientific statement from the American Heart Association. *Circulation*, 2009, 119, 2408–2416.
- [13] Steyerberg, E. W., Vickers, A. J., Cook, N. R., et al.: Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*, 2010, 21, 128–138.
- [14] Förhécz, Z., Gombos, T., Borgulya, G., et al.: Red cell distribution width in heart failure: prediction of clinical events and relationship with markers of ineffective erythropoiesis, inflammation, renal function, and nutritional state. *Am. Heart J.*, 2009, 158, 659–666.
- [15] Jenei, Z. M., Gombos, T., Förhécz, Z., et al.: Elevated extracellular HSP70 (HSPA1A) level as an independent prognostic marker of mortality in patients with heart failure. *Cell Stress Chaperones*, 2013 Apr 7. [Epub ahead of print]
- [16] Kundu, S., Aulchenko, Y. S., van Duijn, C. M., et al.: PredictABEL: an R package for the assessment of risk prediction models. *Eur. J. Epidemiol.*, 2011, 26, 261–264.
- [17] Robin, X., Turck, N., Hainard, A., et al.: pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, 2011, 12, 77.
- [18] DeLong, E. R., DeLong, D. M., Clarke-Pearson, D. L.: Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, 1988, 44, 837–845.
- [19] Cook, N. R., Ridker, P. M.: Advances in measuring the effect of individual predictors of cardiovascular risk: the role of reclassification measures. *Ann. Intern. Med.*, 2009, 150, 795–802.
- [20] Yates, J. F.: External correspondence: Decompositions of the mean probability score. *Organiz. Behav. Human Perform.*, 1982, 30, 132–156.
- [21] Pencina, M. J., D'Agostino, R. B. Sr., Steyerberg, E. W.: Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat. Med.*, 2011, 30, 11–21.
- [22] Hilden, J., Habbema, J. D., Bjerregaard, B.: The measurement of performance in probabilistic diagnosis. II. Trustworthiness of the exact values of the diagnostic probabilities. *Methods Inf. Med.*, 1978, 17, 227–237.

(Jenei Zsigmond Máté dr.,  
Budapest, Kútvölgyi út 4., 1125  
e-mail: jzsig04@gmail.com)

## Állásajánlat!

Orvosokat keresünk német és osztrák kórházakba (minden szakirányban).

**Kiemelkedő fizetés.**

**Információ:**

Tel.: +49/631/3104 2401

Fax.: +49/631/3104 2402

Kapcsolattartó: Frau Koehler, Euro-Matrix GmbH

mail@euro-matrix.de

www.euro-matrix.de