# A teljes emlőbesugárzás pozíciójának előrejelzését szolgáló döntéstámogató statisztikai módszerek

Ph.D. Tézis

Interdiszciplináris Orvostudományok Doktori Iskola

**Rárosi Ferenc**

Szegedi Tudományegyetem, Orvosi Fizikai és Orvosi Informatikai Intézet

Témavezetők:

**Dr. Boda Krisztina, Ph.D.**

**Prof. Dr. Nyári Tibor, Ph.D., D.Sc.**

Szegedi Tudományegyetem, Orvosi Fizikai és Orvosi Informatikai Intézet

**Szeged, 2020**

# 1. Bevezetés

A biostatisztika a matematikai statisztika alkalmazása élettudományi kutatásokban. A biostatisztika módszerei igen nagy területet ölelnek föl a leíró statisztikától a hipotézisvizsgálatokon át a többváltozós statisztikai modellekig. Különös jelentőséggel bírnak a predikciós modellek, amelyek segítenek egy betegséggel (vagy állapottal) kapcsolatos rizikótényezők, vagy prediktorok prediktív erejét föltérképezni és megfelelő prediktorokat „jósló változókat" kiválasztani.

## 1.1. A teljes emlőbesugárzás

A bal oldali teljes emlő sugárterápiájának tervezésekor megfigyelték, hogy bizonyos betegek rizikószerv terhelése (különösen LAD „*left anterior descending coronary artery" „bal elülső koronária leszálló ág"*) függ a besugárzási pozíció megválasztásától (hasi, vagy háti fektetésben történő besugárzás). A kedvezőbb pozíció megválasztása kiválasztható mindkét pozícióban elkészült besugárzási tervek összehasonlítása során, ezt tekinthetjük optimális, a következőkben „gold standard"-nak nevezett eljárásnak. Az eljárás igen sok munkával jár, és szükséges hozzá mindkét pozícióban egy-egy teljes CT sorozat elkészítése. Indokolt tehát olyan statisztikai alapú eljárást keresni, ami bizonyos betegre jellemző adatokat fölhasználva képes jó hatásfokkal előrejelezni a kedvezőbb pozíciót.

## 1.2. Célkitűzések

Interdiszciplináris együttműködésünk céljai az alábbiak voltak:

(1) A függő változó (LAD dóziskülönbség a két pozícióban készített tervek alapján) és a lehetséges prediktorok közötti kapcsolat föltérképezése. A lehetséges prediktorok kiválasztása.

(2) Statisztkai modellen alapuló osztályozási eljárások alkalmazása a kedvezőbb besugárzási pozíció előrejelzésére.

(3) A lehetséges előrejelző (predikciós) modellek jóságának (prediktív erejének) vizsgálata. A különböző predikciós modellek összehasonlítása a döntési görbék új szemléletű módszerével. Szimulációk végzése, hogy megerősítsük a döntési görbe eljárás alkalmazhatóságát a mi modelljeinkre.

(4) A kiválasztott végső modell belső validálása keresztvalidálás segítségével és a modell validálása új betegek adatain (független adatokon).

(5) A prediktorok esetleg mérési hibával terheltek, ezért célunk ezen mérési hibák vizsgálata és különösen annak vizsgálata, hogy ezek a mérési hibák milyen hatással vannak az előrejelzéseinkre.

(6) A LAD dózisok empirikus eloszlásának vizsgálata és Kolmogorov-Szmirnov konfidencia sáv szerkesztése az eloszlásra. Dózisküszöb értékek meghatározása.

Reményeink szerint ez a munka módszereit tekintve esetleg jövőbeni kutatások alapjául is szolgálhat.

Végül, de nem utolsósorban matematikai statisztika módszereinek segítségével hasznos orvosi eszközt szeretnénk nyújtani, ami segíti a besugárzástervezés során a szakemberek munkáját.

# 2. Módszerek

## 2.1. Gyakran használt predikciós modellek

Többváltozós logisztikus regressziós és többváltozós lineáris regressziós modelleket alkalmaztunk, célunk volt az is, hogy könnyen alkalmazható és nem túl bonyolult, de kellően hatásos modellt találjunk.

## 2.2. Mérőszámok a módszerek prediktív erejének jellemzésére

### *Klasszikus mérőszámok*

Ahogy a gyakorlatban minden diagnosztikus tesztnél, úgy minden predikciós modellnél, itt is számítani kell arra, hogy téves előrejelzések előfordulhatnak. Az előrejelzés eredményét a „gold standard” eljárással egybevetve lehetőségünk nyílik a predikciós modell „jóságának”, prediktív erejének jellemzésére. Ezek a jól ismert mérőszámok a szenzitivitás, specificitás, helyesen diagnosztizáltak aránya stb. az egybevetés számszerűsítésének eredményei és a valós pozitív (TP), álpozitív (FP), valós negatív (TN), álnegatív (FN) esetek számain alapulnak.

### *ROC analízis*

A ROC analízist először a Második Világháborúban a rádiójelek vételének jósága jellemzésére használták. Kiderült, hogy a ROC (Receiver Operating Characteristics) analízis igen jól alkalmazható folytonosnak tekinthető prediktor prediktív erejének jellemzésére. Különösen fontos a ROC görbe alatti terület $AUC_{ROC}$. Ez ideális esetben, a tökéletes elválasztás esetében 1, véletlenszerű találgatás esetében 0,5. Fontos továbbá a 95%-os konfidenciaintervallum (95% K-I) a görbe alatti területre.

## *Döntési görbe elemzés (DCA)*

A ROC analízis széleskörben elterjedt módszer folytonos prediktorok prediktív erejének vizsgálatára, azonban a kimenetelek természetes súlyozása nem része a módszernek. A döntési görbe elemzés egy lehetséges megoldást kínál a súlyozás hiányára. A döntési görbe módszer az úgynevezett „nettó nyereség", („net benefit") mennyiségen alapul, ami a döntés klinikai hasznosságát is figyelembe veszi.

## *Egy új mérőszám: a "nettó nyereség"*

Peirce 1884-ben az alábbi mennyiséggel definiálta egy előrejelzés „hasznosságát":

$$B = \frac{p \cdot TP - l \cdot FP}{TP + FP + FN + TN} = \frac{p \cdot TP - l \cdot FP}{N}, \qquad (1)$$

ahol $p$ a valódi pozitív esetben történő előrejelzés hasznossága („*profit*"), míg $l$ az álpozitív döntés (pl. téves riasztás) miatti veszteség („*loss*"), $TP$ a valódi pozitív esetek száma, $FP$ az álpozitív esetek száma, $N$ az összes esetszám. [Peirce, C.S., *The numerical measure of the success of predictions.* Science, 1884. **4**(93): p. 453-4.]

Nettó nyereségnek („net benefit") a följebb definiált „hasznosság" és a „profit" hányadosát értjük:

$$NB = \frac{B}{p} = \frac{TP}{N} - \frac{l}{p}\frac{FP}{N}. \qquad (2)$$

A legtöbb esetben azonban sem $p$, sem $l$ nem ismert mennyiségek (különösen igaz ez az orvos döntések esetében), mégis meglepő módon az $l$ és $p$ arányáról tudunk nyilatkozni. Fontos feltétel az alábbi:

$$\frac{l}{p} := \frac{p_t}{1 - p_t}, \qquad (3)$$

ahol $p_t$ a valószínűségi elvágópont. [Bővebben a feltételről: Vickers, A. J. és Elkin, E. B. (2006). Decision curve analysis: A novel method for evaluating prediction models. Medical Decision Making 26, 565–574.]

Ha ezt a feltételt elfogadjuk, akkor a nettó nyereség már számítható a $p_t$ valószínűségi elvágópont és $TP$ (valódi pozitív esetek száma) és $FP$ (az álpozitív esetek száma) ismeretében:

$$NB = \frac{B}{p} = \frac{TP}{N} - \frac{p_t}{1 - p_t}\frac{FP}{N}. \qquad (4)$$

A „nettó nyereséget" a valószínűségi elvágópont függvényében (y tengelyen a „nettó nyereséget", x tengelyen a lehetséges elvágópont értékeket) ábrázolva kapjuk az ún. „döntési görbét" (decision curve). A döntési görbék módszerét alapvetően olyan modellek értékelésére lehet használni, ahol a modell valószínűséget becsül.

### *Döntési görbék "nem-valószínűségi predikció" esetén*

A döntési görbék módszere valószínűségi előrejelzés esetében használható közvetlenül. Fölmerül azonban a kérdés, hogy folytonos eloszlásúnak tekinthető prediktor esetében hogyan lehet használni ezt a módszert. Természetes módon adódik az ötlet, hogy valószínűségi becslést kell alkalmaznunk és a valószínűségi elvágópontokat fölhasználva már megszerkeszthető a döntési görbe, azonban több lehetőségünk is van arra, hogy a prediktor értékeit a valószínűségi skálára [0,1] képezzük le. Ez fölveti a kérdést, hogy a döntési görbe alakja függhet a leképezéstől, tehát a „link függvény" megválasztásától. A kérdés megválaszolására szimulációkat készítettünk. Az IBM SPSS 24 és R (3.3.1) statisztikai programcsomagokat használtuk.

## 2.3. Teljes emlőbesugárzás hason vagy háton történő pozicionálása

### *Adatbázis, lehetséges prediktorok*

Az első adatbázisunk 83 beteg adatait tartalmazta, ezt követően a módszerek kidolgozásához és a modellek illesztéséhez rendelkezésünkre állt további 55 beteg adata (összesen 138), végül ezeken túl további mintegy 100 beteg adataival rendelkeztünk a prediktorok leolvasási hibájának értékelésekor. (Az etikai engedély száma 185/2012.) A vizsgált prediktorok a következők voltak: testmagasság, testtömeg, testtömeg index (BMI), emlőtérfogat („planning target volume", „PTV"), derék körméret, csípő körméret. Későbbi prediktorok a mezőbe eső szív keresztmetszet „area", a LAD és a mellkas távolsága medián értéke „medián távolság" „median distance". A függő változó a LAD dóziskülönbség volt (háti-hasi pozícióban, ezt a tanuló adathalmazon mindkét pozícióban elkészített besugárzási tervekből tudtuk, a LAD dóziskülönbség előjele meghatározta, hogy melyik a kedvezőbb pozíció). A tervezőrendszerből a háti és hasi pozícióban tervezett LAD dózisok különbségét képezve megkaptuk a LAD dóziskülönbség értékét. Ha ez pozitív, akkor a hasi pozíció a kedvezőbb (hasi fektetésben lenne kisebb a LAD terhelése), ha negatív, akkor a háti pozíció a kedvezőbb LAD terhelés szempontjából. Célunk volt a tervezőrendszerből ismert dóziskülönbséget, vagy legalábbis a kedvezőbb pozíciót minél jobb hatásfokkal becsülni egyszerű módszerekkel, ezáltal olyan módszert kidolgozni, ami a további betegek esetében lehetővé teszi a jó hatásfokú előrejelzést, és könnyíti a besugárzástervezés menetét. A LAD dózisa mellett vizsgálataink egy része kiterjedt a 25 Gy-nél nagyobb dózist kapó relatív szívtérfogatra ($V_{25Gy}$ %) is, ezt másodlagos kimeneti változónak tekintettük. A részletes módszertant és részletes eredményeket a LAD dóziskülönbségre dolgoztuk ki, de a $V_{25Gy}$ -re vonatkozó eredményeinkről is olvashatunk a tézis alapjául szolgáló publikációinkban.

### *Keresztvalidálás*

Jól ismert tény, hogy ha ugyanazon a mintán vizsgáljuk egy predikciós modell döntésének jóságát, mint amin a modell együtthatóit becsültük, akkor indokolatlanul optimista eredményt kapunk. Ezért célszerű a rendelkezésünkre álló adatokat két halmazra bontani, az egyik a „tanulóhalmaz" („training set"), amin az együtthatók becslése történik, a másik a „teszt halmaz" („test set"), amin a becslések jóságát ellenőrizzük. Vizsgálatunkban ezerszeres véletlenszerű keresztvalidálást alkalmaztunk 3 SPSS makró segítségével. A regressziós együtthatók becslése mindig az adatok véletlenszerűen kiválasztott 70%-án történt, majd a maradék 30%-on kiértékeltük a modell által szolgáltatott előrejelzés jóságát.

### *Bland-Altman módszer az egyezés mérésére*

A prediktorok esetlegesen leolvasási hibával terheltek lehetnek. Vizsgálataink harmadik fázisában (amikor 100, az előzőektől független beteg adatait használtuk) bizonyos prediktor értékeket („area" és „medián távolság") egyetlen CT szelet alapján is becsültük, illetve egy egy CT sorozat alapján is becsültük ugyanazon betegek esetében is. A CT sorozatból nyert „area" és „medián távolság" prediktor értékeket tekinthetjük referenciának, mert a CT sorozat fölhasználása több információt jelent, mint egyetlen CT szelet fölhasználása. Mivel 100 beteg adatain mindkét eljárással megvizsgáltuk ezen értékeket, így lehetőségünk nyílt ezek egybevetésére. Ehhez az egybevetéshez az orvosi statisztikában méltán ismert és elismert Bland-Altman módszert vettük igénybe.

### *A Kolmogorov-Szmirnov konfidencia sáv*

Igen fontosnak találtuk, hogy ne csak a LAD dóziskülönbségeket vegyük vizsgálat alá, hanem az „abszolút" LAD dózisok eloszlását is. Például 2 Gy dóziskülönbség lehet az eredménye 4 Gy és 2 Gy dózisoknak, de lehet az eredménye 22 Gy és 20 Gy dózisoknak is. A LAD dózisa a kedvezőbb pozícióban természetesen a hasi és háti pozícióban tervezett értékek minimuma. Hogy megvizsgáljuk mi az elérhető realitás, a kedvezőbb pozícióban megvizsgáltuk a dózisok empirikus eloszlását és 95%-os Kolmogorov-Szmirnov konfidencia-sávot szerkesztettünk az empirikus eloszlásfüggvényre. Percentilis becslés segítségével dózisküszöböket határoztunk meg, (amely tervezett dózist meghaladva egyéb intervenciót tartunk szükségesnek).

# 3. Eredmények

## 3.1. A prediktorok működése

Önmagában egyik prediktor sem volt elegendően jó megbízhatóságú előrejelzéshez. A ROC analízis eredménye, hogy a görbe alatti területek ($AUC_{ROC}$) a következőképpen alakultak: a legjobb prediktornak az „area” adódott $AUC_{ROC}$=0,868 (95% K-I: 0,791; 0,944). „medián távolság” esetében $AUC_{ROC}$=0,787 (95% K-I: 0,690; 0,884). A legrosszabb a testtömeg index „BMI” $AUC_{ROC}$=0,740 (95% K-I: 0,630; 0,850).

## 3.2. Predikciós modellek

### *Logisztikus regresszión alapuló modellek*

Mivel önmagában egyik prediktor sem felelt meg a céljainknak, ezért többváltozós statisztikai modellre volt szükség. A multikollinearitás kiküszöbölése érdekében hierarchikus klaszterezés alá vontuk a „prediktor jelölteket” és gyengén korreláló változókat használtunk föl a modell építéséhez. Először a többváltozós logisztikus regresszión alapuló modelleket vizsgáltuk. „Befelé léptető” (forward) és „kiléptető” (backward) likelihood-hányados modell szelekciós eljárásokat használtunk.

A backward likelihood-hányados kritérium alapján választott logisztikus regressziós modell a „főhatás-modell” volt: *area+BMI+median distance*.

A forward likelihood-hányados kritérium alapján választott logisztikus modell ezzel szemben egy ún. „nem hierarchikus modell” volt. Ez azt jelentette, hogy a prediktorok interakciója (szorzata) szerepelt a modellben, de a főhatások maguk nem: *area*BMI+area*median distance*. A nem hierarchikus modellek együtthatóinak jelentését igen nehéz lenne interpretálni, azonban ez nem célunk. Célunk a gyakorlati fölhasználás az előrejelzésre nem pedig elvi mechanizmusokat vizsgálni.

Mindkettő logisztikus modellre igaz, hogy az együtthatói szignifikánsan különböznek nullától még 1%-os szinten is. Lehetséges a logisztikus regresszió által adott valószínűségi előrejelzéseket is ROC analízissel vizsgálni. Ez az analízis föltárta, hogy mindkettő logisztikus modell jól használható előrejelzésre. A „főhatás-modell” esetében $AUC_{ROC}$=0,906 adódott (95% K-I: 0,854; 0,959), míg a „nem hierarchikus modell” esetében $AUC_{ROC}$=0,900 (95% K-I: 0,848; 0,953).

### *A lineáris regresszión alapuló modell*

Bár a logisztikus regresszión alapuló modellek prediktív ereje viszonylag jónak mutatkozott, hiányossága a logisztikus regressziós modellek alkalmazásának, hogy a LAD

dóziskülönbségnek csak az előjelét veszi figyelembe (ugyanis melyik a kedvezőbb pozíció), a dóziskülönbség értékét azonban nem. Ez azt jelenti, hogy nem használjuk ki a rendelkezésre álló információt és nem történik súlyozás a dóziskülönbség szerint. Ha a függő változót folytonos eloszlásúnak tekintjük, akkor regressziós problémával állunk szemben. Ebben az esetben a nagyobb dóziskülönbségek nagyobb súllyal kerülnek a modellbe és az eljárás a dóziskülönbséget magát is becsüli, nem csak a kedvezőbb pozíciót.

A többszörös lineáris regressziós modellre is igaz, hogy minden együtthatója szignifikáns 1%-os szinten is, leszámítva a konstanst, amire p=0,012 teszt p-érték adódott. A modell prediktív ereje is hasonlóan jónak mutatkozott a kettő logisztikus regressziós modelléhez, $AUC_{ROC}$=0,903-at kaptunk (95% K-I: 0,850; 0,957).

### *Illeszkedési jóság és modell diagnosztika*

A többszörös korrelációs együtthatóra *R=0,754*-et kaptunk, az igazított $R^2_{adj}$ értékére 0,560-at.

Bár a fenti eredmények biztatóak voltak, igen fontos a lineáris regressziós modell feltételei teljesülését megvizsgálni. A reziduumok normalitását Q-Q ábrával és Shapiro-Wilk próbával vizsgálva is megnyugtató eredményre jutottunk, a Shapiro-Wilk próba p-értéke: *p=0,592*. Grafikusan meggyőződtünk a reziduumok állandóságáról (a reziduumokat a lineáris regresszió által előrejelzett értékek függvényében kirajzolva nem mutatkozott trend) és arról, hogy a reziduumok az idővel nem mutatnak tendenciát (a beteg sorszámát és a reziduumokat ábrázolva szintén nem mutatkozott trend). A reziduumok függetlenségét Durbin-Watson próbával vizsgálva megnyugtató eredményre jutottunk, a Durbin-Watson próbastatisztika értéke (*d=1.847*) az elfogadási tartományba esett.

### *Modell validálás*

Az ezerszeres keresztvalidálás segítségével nem csak a lineáris regressziós modell előrejelzésének szenzitivitását és specificitását, hanem a tévesen előrejelzett esetekben a dózisok eloszlását is figyelembe tudtuk venni. (A szenzitivitás és specificitás értelmezése kapcsán fontos megjegyezni, hogy a háti pozíciót tekintettük pozitív kimenetelnek.) Ennek eredményeit az alábbi táblázat tartalmazza. Optimális elvágópontnak a 0,6 Gy-t tartottuk, ebben az esetben a szenzitivitásra 80,7%, a specificitásra 87,5% jut a LAD dóziskülönbség szempontjából az átlagos tévedés kiegyensúlyozott és mértéke ±1,7 Gy.

| Elvágópont | Szenzitivitás(%) | Specificitás(%) | Tévedés mértéke, döntés: hasi pozíció (Gy, átlag±SD) | Tévedés mértéke, döntés: háti pozíció (Gy, átlag±SD) |
|---|---|---|---|---|
| -0,6 | 66,6 | 91,1 | 2,5±3,9 | -0,7±1,0 |
| -0,3 | 70,8 | 90,7 | 2,6±3,6 | -0,8±1,1 |
| 0 | 74,4 | 90,0 | 2,4±3,4 | -0,9±1,3 |
| 0,3 | 77,7 | 88,9 | 2,1±3,0 | -1,2±1,6 |
| 0,6 | 80,7 | 87,5 | 1,7±2,6 | -1,7±1,9 |
| 0,9 | 83,4 | 86,0 | 1,5±2,4 | -2,0±2,2 |
| 1,2 | 85,4 | 83,6 | 1,1±2,3 | -2,3±2,8 |
| 1,5 | 86,5 | 81,7 | 1,1±2,2 | -3,0±3,7 |
| 1,8 | 86,8 | 79,9 | 1,3±2,3 | -3,5±4,2 |

1. Táblázat. *A lineáris regressziós modell ezerszeres keresztvalidálásának eredményei. (a szenzitivitás és specificitás értékeket úgy kell érteni, hogy a háti pozíciót tekintjük „pozitivnak”, a hasi pozíciót „negatívnak”)*

## 3.3. Néhány eredmény a döntési görbék alkalmazhatósága kapcsán, nem valószínűségi predikció esetén

Célunk volt a modellek összehasonlítása a döntési görbék módszerével is. Ehhez azonban szükségünk volt arra, hogy a lineáris regresszió által becsült dóziskülönbség értékekből valószínűségi becslést készítsünk. Általánosságban is elmondható, hogy ha a döntési görbéket nem valószínűségi predikció esetén szeretnénk alkalmazni, szembe kerülünk azzal a nehézséggel, hogy miként transzformáljuk a folytonos változó értékeit a valószínűségek [0,1] intervallumába.

Bár kézenfekvő választás a logisztikus regresszió által becsült valószínűségeket fölhasználni a döntési görbe szerkesztéséhez, vannak más indokolható lehetőségek is. Ilyenek az empirikus eloszlásból származó valószínűségek használata, az inverz logit transzformáció, vagy a probit-link függvény használata.

Röviden összefoglalva szimulációink megerősítették, hogy előnyös a logisztikus regresszió által becsült valószínűségeket használni a döntési görbék megszerkesztéséhez. Különösen igaz ez olyan esetekben, amikor a „nem valószínűségi előrejelzés” során kapott érték ferde eloszlást követ.

## 3.4. Modelljeink összehasonlítása a döntési görbék módszerével

A szimulációk eredményei lehetővé tették, hogy a lineáris regresszión alapuló modell értékelésénél is használhassuk a döntési görbék módszerét. A mindhárom többváltozós statisztikai modell (a logisztikus regressziós „főhatás modell”, a logisztikus regressziós „nem hierarchikus modell” és a lineáris regresszión alapuló modell) igen hasonló és magas „nettó

nyereség” („net benefit”) értékeket eredményezett, ami azt mutatja, hogy mindhárom többváltozós modell előnyösen alkalmazható döntéselméleti szempontból.

## 3.5. Prediktor mérési hibák és következményeik független betegek adatain vizsgálva

### *Prediktor mérési hibák és következményeik*

Száz, az előzőektől különböző beteg adatait vizsgálva mind CT sorozatból, mind egyetlen kiválasztott CT szelet alapján megvizsgáltuk az „area” és a „medián távolság” prediktor értékeket. A CT sorozat alapján kapott prediktor értékeket tekinthetjük a pontosabbnak, mert több információt használ föl, mint az „egyszeletes” értékek. A CT szelet kiválasztásának módját az alábbi cikkben publikáltuk. [III]

A Bland-Altman módszer föltárta, hogy az „area” tekintetében a 95%-os egyetértési határok az egyszeletes és a CT sorozat alapján történő mérést egybevetve: -5.624 $cm^2$-től 3.902 $cm^2$-ig terjednek, „medián távolság” esetében -0.829 cm-től 0.783 cm-ig terjednek. Ennél lényegesebb kérdés, hogy a lineáris regresszió egyenletébe behelyettesítve az értékeket mekkora különbségeket kapunk a modellünk által becsült dózisokat tekintve az egyszeletes, illetve a sorozatból származó prediktor értékeket fölhasználva. A 95%-os egyezési határok: -6.65 Gy-től 7.82 Gy-ig terjednek. Bár ezek széles intervallumot határoznak meg, a mi célunk végső soron a kedvezőbb pozíció előrejelzése és meggyőződtünk róla, hogy az egyszeletes prediktor értékek használata nem vezet érdemi hiba halmozódáshoz. Az ezerszeres keresztvalidálás eredménye volt, hogy a 0,6 Gy elvágópont mellett a szenzitivitás 80,7%, a specificitás 87,5%. A 100 újabb beteg adatain az egyszeletes prediktor értékeket a lineáris regresszió egyenletébe helyettesítve az keresztvalidálással rendkívül konzisztens eredményt kaptunk, a szenzitivitásra 82.8%-ot, a specificitásra 87.3%-ot.

### *A dózis értékek empirikus eloszlása Kolmogorov-Szmirnov konfidencia sáv, dózisküszöbök*

Amint azt említettük az abszolút dózisok vizsgálata is rendkívül fontos, hiszen elképzelhető pl. kis dóziskülönbség egyébként nagy abszolút dózis értékek esetében is. A kedvezőbb pozícióban jellemző LAD dózis értékeket a tervező rendszer segítségével ismertük meg. Megvizsgáltuk a kedvezőbb pozícióban levő LAD dózis értékek empirikus eloszlását. 90%-os percentilis becsléssel határoztuk meg azon dózisküszöböket, amelyek meghaladása esetében egyéb intervenciót tartottunk szükségesnek. A 90%-os percentilis értékekre 12,5 Gy és 12,9 Gy adódott a háti illetve a hasi kedvezőbb pozíció esetében. A 95%-os Kolmogorov-Szmirnov

konfidencia-sáv megerősítette, hogy rossz esetben is csak a betegek legfeljebb 30%-a esetében adódik ennél nagyobb LAD dózis érték a kedvezőbb pozícióban.

### *A modell külső adatokon történő vizsgálata*

A fenti adatok mindegyike az SzTE ÁOK Onkoterápiás Klinika betegei adataira támaszkodott, ami fölvetheti a külső validálás (másik ellátóhely betegeinek adatain) iránti igényt. Biztató eredmény, hogy a Liege-i Sugárterápiás Központban 28 betegen megvizsgáltuk a lineáris regresszión alapuló módszer hatékonyságát, és az 24 esetben helyes döntésre vezetett, a helyes előrejelzések arányára 85,7% adódott.

## 4. Diszkusszió

Interdiszciplináris együttműködésünk keretében terápiás predikciós modelleket dolgoztunk ki. Ezek a modellek az orvosi statisztikában igen megalapozott logisztikus és lineáris regresszió módszerre támaszkodtak.

Részletesen bemutattuk a modellek hatékonyságának, prediktív erejének vizsgálatát. Szimulációink megerősítették, hogy a döntési görbék új szemléletű módszerét hogyan lehet használni „nem valószínűségi" előrejelzés (például regresszió által becsült várható értékek) esetén. A „belső validálás" céljait a széleskörűen elfogadott keresztvalidálás szolgálta. Az „időbeli validálás" és a prediktor leolvasási hibák vizsgálata újabb 100 beteg adatain történt. A két híres orosz matematikus Kolmogorov és Szmirnov, rendkívüli jelentőségű és nagy idézettségű módszerét (konfidencia-sáv az empirikus eloszlásra) jó eredménnyel használtuk a rizikószerv dózis terhelés értékek eloszlásának vizsgálatára és a dózisküszöb értékek meghatározására.

A prediktorok egyike sem bizonyult kielégítőnek önmagában, lényegesen hatékonyabb volt a többváltozós statisztikai modellek mindegyike. A lineáris regresszión alapuló modellt tekintettük a legmegfelelőbbnek, mert az a rizikószerv terhelés értékek szerint is súlyoz és minden egyéb szempontból összevethető a logisztikus regressziós modellek eredményeivel. A döntési görbék módszere föltárta, hogy mind a logisztikus, mind a lineáris regressziós modellek magas „nettó nyereség" („net benefit") értékeket eredményeznek a valószínűségi elvágó pontok széles sávjában. A lineáris regressziós modell hatékonysága igen stabilnak mutatkozott a validálás lépései során. A Kolmogorov-Szmirnov konfidencia sáv alapján kimondható, hogy 95% valószínűséggel legfeljebb a betegek 30%-a esetében haladja meg a LAD dózis értéke az általunk meghatározott dózisküszöböt.

Vizsgálatunk erőssége a nagy mintaelemszám és a többváltozós szemlélet. Bár a ROC analízis elterjedt a besugárzástervezésben is, de a döntési görbék módszerének alkalmazásával ezen a területen még nem találkoztunk.

Szólnunk kell a limitációkról is. A lineáris regressziós modell egyszerűen elsőfokú összefüggést tételez föl a prediktor értékek és a függő változó között (másképpen a három prediktor és a LAD dóziskülönbség által meghatározott négydimenziós térben egy háromdimenziós hipersíkot illesztettünk), azonban a pontdiagramok alapján nem találtunk más jellegű összefüggést, ami alkalmasabb lett volna a becslésre.

Munkánk elején a prediktor értékek CT sorozaton alapultak. Még ezek az értékek is lehetnek hibával terheltek (pl. leolvasási hiba). Ezt azonban nem volt lehetőségünk vizsgálni, mert nem történtek erre vonatkozó ismételt mérések. Előbbivel összefüggésben bírálat érheti a hagyományos legkisebb négyzetes becslés (OLS „ordinary least squares estimation”) alkalmazását, azonban a reziduumok vizsgálata nem mutatott összefüggést a függő változó (LAD dóziskülönbség) varianciája és a regresszió által becsült értékek között. A módszer feltételeit részletekbe menően vizsgáltuk, és megnyugtató eredményeket kaptunk; célunk egy predikcióra használható modell létrehozása volt, nem olyan modellé, ami a háttérben levő mechanizmusokat magyarázza.

Természetesen a lineáris regressziós modell prediktív erejének megvannak a maga korlátai, ezek jellemzően 80,7% körüli szenzitivitás és 87,5% körüli specificitás, azonban a modell a maga korlátaival éveken át stabilan és hasznosan működött a klinikai gyakorlatban. Bízunk benne, hogy munkánk segíti az egyénre szabott terápia (helyesen megválasztott kezelési pozíció) céljait és a rizikószervek sugárterhelésének csökkentését.

# 5. Következtetések

A döntési görbék módszere hatékony eszköznek bizonyult predikciós modelljeink összehasonlítása céljából. Lineáris regressziós modellünket hatékony eszköznek tekintjük, ami segítheti az orvosok és a fizikusok munkáját a mindennapi klinikai gyakorlatban az előnyös fektetési pozíció előrejelzésével.

- Eredményeink föltárták az összefüggést a prediktorok: a testtömeg index (BMI), „medián távolság” („median distance”), mezőbe eső szív keresztmetszet („area”) és a függő változó (LAD dóziskülönbség) között.
- A többszörös lineáris regresszión alapuló modell hatékonynak bizonyult a bal oldali emlőrákos betegek besugárzási pozíciójának előrejelzése szempontjából.
- Megmutattuk, hogy a döntési görbe módszer jól alkalmazható „nem valószínűségi predikció” esetében is. Szimulációs eredményeink megerősítik a logisztikus regresszió által becsült valószínűségek használatának helyességét a „nettó nyereség” („net benefit”) számolásához ilyen esetben.
- A többszörös lineáris regressziós modell prediktív erejének vizsgálata során a döntési görbék új szemléletű módszere jól használhatónak bizonyult.
- A döntési görbék elemzése és a véletlenszerű keresztvalidálás megmutatják a többszörös lineáris regressziós modell használhatóságát.
- A többszörös lineáris regresszión alapuló modellt használják (az általunk megállapított dózis feltételek figyelembevételével) az SzTE ÁOK Onkoterápiás Klinika kutatói és orvosai [III]. A többszörös lineáris regresszión alapuló modell alkalmas a besugárzási pozíció előrejelzésére.

# 6. Köszönetnyilvánítás

Mindenekelőtt köszönettel tartozom Dr. Boda Krisztinának és Prof. Dr. Nyári Tibornak, témavezetőimnek a lelkiismeretes segítő munkájukért, gondos tanácsaikért, a szakmai támogatásukért.

Rendkívül hálás vagyok Kahán Zsuzsanna Professzor Asszonynak, az SzTE ÁOK Onkoterápiás Klinika volt igazgatójának, aki töretlen optimizmusával és folyamatos bátorításával minden nehézségen átsegített és gondos szakmai tanácsaival mindig segített.

Köszönetet mondok Dr. Kószó Renáta Lillának és Dr. Varga Zoltánnak (SzTE ÁOK Onkoterápiás Klinika) gondos munkájukért és a lehetőségért, hogy velük dolgozhattam.

Köszönetet mondok minden kedves társszerzőnek, akivel együtt dolgozhattam.

Szeretném kifejezni köszönetemet Prof. Dr. Bari Ferencnek és Prof. Dr. Peták Ferencnek az SzTE ÁOK Orvosi Fizikai és Orvosi Informatikai Intézet volt és jelenlegi vezetőjének töretlen biztatásukért és azért, hogy biztosították számomra a munka elkészüléséhez szükséges körülményeket.

Köszönöm az Orvosi Fizikai és Orvosi Informatikai Intézet dolgozóinak, kedves munkatársaimnak a segítőkészségüket és a barátságos munkahelyi légkört.

Külön köszönet illeti Mező Zsuzsa kolléganőmet, aki a magyar nyelvű tézisfüzet nyelvhelyességét nagy gondossággal ellenőrizte.

Szeretnék köszönetet mondani volt Tanáraimnak, Professzoraimnak, Mentoraimnak lelkiismeretes pedagógiai és szakmai munkájukért.

Hálásan köszönöm barátaimnak a folyamatos erkölcsi és emberi támogatást.

Ezzel a munkával szeretném megköszönni Szüleimnek azt, hogy támogatták tanulmányaimat úgy erkölcsi, mint lelki szempontból. Támogatásuk és töretlen optimizmusuk nélkül ez a munka nem jöhetett volna létre.

A kutatás részben az EFOP-3.6.1-16-2016-00008 pályázat támogatásával valósult meg.

# Az értekezés alapjául szolgáló közlemények

I. Rárosi, F., Boda, K., Kahán, Z., and Varga, Z. (2019). Decision curve analysis apropos of choice of preferable treatment positioning during breast irradiation. BMC Medical Informatics and Decision Making 19(1):204. **IF: 2,067**

II. Varga, Z., Cserháti, A., Rárosi, F., Boda, K., Gulyás, G., Együd, Z., and Kahán, Z. (2014). Individualized positioning for maximum heart protection during breast irradiation. Acta Oncologica 53, 58–64. **IF: 2,997**

III. Kahán, Z., Rárosi, F.,Gaál, Sz., Cserháti, A., Boda, K., Darázs, B., Kószó, R., Lakosi ,F., Gulybán, Á., and Varga, Z. (2018). A simple clinical method for predicting the benefit of prone vs. supine positioning in reducing heart exposure during left breast radiotherapy. Radiotherapy and Oncology 126(3):487-492. **IF: 4,942**