

A végpont és meghatározása

Ferenci Tamás
tamas.ferenci@medstat.hu

Utoljára frissítve: 2024. november 19.

- Mi az, hogy végpont?
- Elsődleges vs. másodlagos
 - „Ami fontos” (klinikailag)
 - Amire a kísérlet ereje tervezve van!
- Mivel írható le? (feltételezzünk az egyszerűség kedvéért kétkarú vizsgálatot)

A végpontok jellegei: bináris (dichotóm) végpont

Bináris – vagy dichotóm – végpont (pl. meghalt-e): egy arányt kapunk minden csoportban (p)

- A „természetes” metrikák:
 - Relatív rizikó (relative risk): $RR = p_T / p_C$ (néha: $RRR = 1 - RR$, pl. a 0,7-es RR-re azt is lehet mondani, hogy 30%-os csökkenés)
 - Abszolút rizikó-csökkenés (absolute risk reduction): $ARR = p_T - p_C$ (néha hívják risk difference-nek is, RD)
- Ez egyszerűsítés volt (de nekünk most elég lesz):
 - az utánkövetési idő hossza – ami eltérhet páciensenként – fontos
 - ha okspecifikus (nem all-cause) mortalitási végpontot használunk, vagy nem mortalitási végpontot használunk, a versengő kockázatok fontosak
- A metrikák általában is besorolhatóak „abszolút” vagy „relatív” típusúként, más jellegű végpontokra is

Bináris végpont egy problémája

- Még a legegyszerűbb esetekben is dilemmákhoz vezethet, egy példa: a kontrollcsoportban 2% kapott infarktust, a sztatinnal kezelt csoportban 1%, akkor most...
 - A fantasztikus gyógyszerünk 50%-kal csökkenti a kockázatot!
 - A fantasztikus gyógyszerünk nélkül 100%-kal nagyobb az infarktus-kockázatod!
 - 100 emberből 98-at *feleslegesen* kezelünk, mert amúgy se kapna infarktust, 1-et *hiába* kezelünk, mert gyógyszerrel együtt is infarktust kap, és csak 1 az, akinél elérünk valamit (viszont közben mind a 100-at kitesszük a mellékhatás-kockázatoknak, mind a 100-zal kifizettjük stb.)!
- Ugye mennyire máshogy hangzik? Pedig csak osztani kell tudni, hogy lássuk: ez a három *igazából ugyanaz!*

Cook, Richard J., David L. Sackett. The number needed to treat: a clinically useful measure of treatment effect. *BMJ*. 310.6977 (1995): 452-454. Malenka DJ, Baron JA, Johansen S, et al. The framing effect of relative and absolute risk. *J Gen Intern Med*. 1993 Oct;8(10):543-8. Gigerenzer G, Wegwarth O, Feufel M. Misleading communication of risk. *BMJ*. 2010 Oct 12;341:c4830.

A „keretezési” hatás

- Malenka (1993): 10% eséllyel meghalsz kezelés nélkül, az A gyógyszer 80%-kal csökkenti ezt a kockázatot, a B gyógyszerrel 100 embert kezelve 8-at megment, melyiket választanád? → A-t 56,8%, B-t 14,7%, mindegy 15,5%, nem tudom 13,0% (a kutatásban betegek vettek részt)
- Bucher (1994): orvosok egy része relatív, más része abszolút formában kapta meg ugyanazon kutatás eredményeit; teljesen máshogy értékelték a gyógyszer jóságát
- Schwartz (1997): az abszolút és relatív információ megértése szorosan összefügg a matematikai alapkészségekkel
- Gigerenzer (2008): „100%-kal növeli a fogamzásgátló a trombóziskockázatot” (1 per 7000 kockázatnövekedés, csak akkora pánik lett belőle Angliában 1995-ben, hogy 13 ezer többlet-terhességmegszakítást hajtottak végre...)

A „keretezési” hatás

- Gigerenzer (2008): „a mammográfia 25%-kal csökkenti az emlőrák-mortalitást” → laikusok több mint negyede szerint ez azt jelenti, hogy minden 1000 nőből 250-nel kevesebb hal meg emlőrákban, még nőgyógyász szakorvosok 15%-a szerint is ez a helyzet

Malenka DJ, Baron JA, Johansen S, et al. The framing effect of relative and absolute risk. *J Gen Intern Med.* 1993 Oct;8(10):543-8. Bucher HC, Weinbacher M, Gyr K. Influence of method of reporting study results on decision of physicians to prescribe drugs to lower cholesterol concentration. *BMJ.* 1994 Sep 24;309(6957):761-4. Moynihan R, Bero L, Ross-Degnan D, et al. Coverage by the news media of the benefits and risks of medications. *N Engl J Med.* 2000 Jun 1;342(22):1645-50. Mason D, Prevost AT, Sutton S. Perceptions of absolute versus relative differences between personal and comparison health risk. *Health Psychol.* 2008 Jan;27(1):87-92. Schwartz LM, Woloshin S, Black WC, et al. The role of numeracy in understanding the benefit of screening mammography. *Ann Intern Med.* 1997 Dec 1;127(11):966-72. Gigerenzer G, Wegwarth O, Feufel M. Misleading communication of risk. *BMJ.* 2010 Oct 12;341:c4830. Gigerenzer G, Gaissmaier W, Kurz-Milcke E, et al. Helping Doctors and Patients Make Sense of Health Statistics. *Psychol Sci Public Interest.* 2007 Nov;8(2):53-96.

- Az előző gondolat elvezet minket az NNT ötletéhez (number need to treat)
- $NNT = \frac{1}{ARR}$
- Hány beteget kell *átlagosan* kezelni egy végpont megelőzéséhez (ez fontos, véletlenség is van, nyilván nem arról van szó, hogy *pontosan* 1-gyel kevesebb végpont lesz NNT fő beteget kezelve)
 - Fix mintanagyság értelmezés: NNT fő kezelt és NNT fő kontroll beteg esetén a kezelt csoportban várhatóan 1-gyel kevesebb végpont lesz (értsd: a végpontok száma és így annak különbsége véletlen, ennek várható értéke 1)
 - Véletlen mintanagyság értelmezés: ez passzol jobban a fenti mondathoz, de zűrösebb, hogy egyáltalán mit értünk alatta

Grieve, A. P. (2003). The number needed to treat: a useful clinical measure or a case of the Emperor's new clothes?. *Pharmaceutical Statistics: The Journal of Applied Statistics in the Pharmaceutical Industry*, 2(2), 87-102.

- A kulcskérdés: ez jelenik meg a kockázat-haszon mérlegben! – a kockázatokat az NNT-vel kell felszorozni
- Már most észrevehető, hogy az NNT nem csak a gyógyszeren múlik, tehát az NNT nem azért volt ennyire magas itt, mert a gyógyszer „rossz”; a dolog a p_C -n, a baseline rizikón (=kontrollcsoport rizikója) is múlik
- Megjegyzés: az NNT statisztikai értelemben elég rossz mutató (torzított, különösen zűrös a CI), de most a koncepció a fontos, és abban nagyon hasznos
- Összefoglalva: az abszolút különbség a releváns, mert mutatja a klinikai hasznot...
- ...a relatív meg a gyógyszergyárak marketing-részlegeinek fontos, hogy feltupírozott számokat tudjanak mutatni??

Grieve, A. P. (2003). The number needed to treat: a useful clinical measure or a case of the Emperor's new clothes?. *Pharmaceutical Statistics: The Journal of Applied Statistics in the Pharmaceutical Industry*, 2(2), 87-102.

- Esély (odds): „valószínűség osztva 1 mínusz valószínűséggel” avagy „eseményt átélők száma osztva az eseményt át nem élők számával” (pl. megbetegedők száma osztva a meg nem betegedők számával – nem az alanyok számával):

$$\text{odds} = \frac{p}{1 - p}$$

- Furcsának hathat, pedig a sportfogadásból ismert a szó: tényleg ugyanarról van szó
- Odds-ból is számolható valószínűség ($p = \frac{\text{odds}}{1 + \text{odds}}$), a kettő között egy-egy megfeleltetés van
- Az, hogy „3 az 1-hez” adják a hazai csapat győzelmét (3 az 1-hez az oddsa) ugyanaz, mint hogy 25% a valószínűsége
- És fordítva: ahelyett, hogy „10% holnap az eső valószínűsége” nyugodtan mondhatná azt az időjárásjelentés, hogy „9 az 1-hez az eső oddsa”
- Mindkettő pontosan ugyanúgy alkalmas valószínűségi információ közlésére, tehát teljes mértékben csak konvenció kérdése, hogy az időjárásjelentésnél ezt szoktuk meg, a sportfogadásnál meg azt

- A két esély hányadosa:

$$OR = \frac{p_T / (1 - p_T)}{p_C / (1 - p_C)}$$

- Tehát teljesen ugyanaz, mint a relatív rizikó, csak „időjárásjelentős” szóhasználat helyett „sportfogadó” szóhasználattal kifejezve a valószínűséget
- Miért jó (azon túl, hogy bizonyos statisztikai modellek ezt szolgáltatják, és hogy eset-kontroll elrendezésnél csak ezt lehet kiszámítani)?
- Az RR nem szimmetrikus, sem arra nézve, hogy melyik irányból nézzük az adatainkat, sem arra nézve, hogy mi a kimenet

Az RR aszimmetriája – hogy nézünk rá az adatokra

Ekcéma és szénanátha kapcsolata:

	Szénanáthás	Nem szénanáthás	Összesen
Ekcémás	141	420	561
Nem ekcémás	928	13525	14453
Összesen	1069	13945	15522

- Szénanátha relatív rizikója ekcéma szerint (ha ekcémás vagy, hányszorosára nő a szénanátha valószínűsége) $\frac{141/561}{928/14453} = 3,91$, ekcéma relatív rizikója szénanátha szerint (ha szénanáthás vagy, hányszorosára nő az ekcéma valószínűsége) $\frac{141/1069}{420/13945} = 4,38$
- Viszont: szénanátha esélyhányadosa ekcéma szerint (ha ekcémás vagy, hányszorosára nő a szénanátha esélye) $\frac{141/420}{928/13525} = 4,89$, ekcéma esélyhányadosa szénanátha szerint (ha szénanáthás vagy, hányszorosára nő az ekcéma esélye) $\frac{141/928}{420/13525} = 4,89!$
- ...miközben a két betegség ebből a szempontból nyilván teljesen szimmetrikus, egyik sem kitüntetett, csak a kapcsolatukra vagyunk kíváncsiak

Bland JM, Altman DG. Statistics notes. The odds ratio. BMJ. 2000 May 27;320(7247):1468.

Az RR aszimmetriája – mit nevezünk kimenetnek

Tekintsük a következő egyszerű példát:

	Meghalt	Túlélt
Kezelt	25	75
Kontroll	50	50

- Ha a halálozás rizikóját nézzük: $RR = \frac{25/100}{50/100} = 0,5$, de ha a túlélés „rizikóját”:
 $RR = \frac{75/100}{50/100} = 1,5$
- Halálozás esélyhányadosa $OR = \frac{25/75}{50/50} = 1/3$, a túlélésé: $OR = \frac{75/25}{50/50} = 3$
- Tehát az esélyhányados szimmetrikus (ha megfordítjuk a kimenetet, megfordul – reciproka lesz – az OR), de a relatív rizikó *nem*: $\frac{1}{0,5} = 2 \neq 1,5$!
- Nyilván teljesen mindegy, hogy a halálozást vagy a túlélést vesszük kimenetnek, hiszen az egyik úgyis meghatározza a másikat, de a relatív rizikó *mégis* függ ettől
- Nagyon zavaró módon *nem ugyanazt* kapom, ha azt számolom, hogy hányan haltak meg és ha azt, hogy hányan éltek túl (noha ez nyilván teljesen mindegy kellene legyen)!

Cummings P. The relative merits of risk ratios and odds ratios. Arch Pediatr Adolesc Med. 2009 May;163(5):438-45.

- Az OR nehezebben értelmezhető
 - „The only way, we are told, that physicians can understand probabilities: odds being a difficult concept only comprehensible to statisticians, bookies, punters and readers of the sports pages of popular newspapers.” (Stephen Senn)
- Az OR konstansabb (homogénebb) lehet
 - Egy 10%-os baseline rizikójú populáción valamilyen káros tényező hatása $RR = 3$ – mi történik, ha ez a tényező egy 40%-os baseline rizikójú populációban hat?!
 - Az 1-nél nagyobb RR-ek matematikailag lehetetlen, hogy mindig érvényesek tudjanak lenni! (Ellenérv: $RR < 1$ -nél nincs ilyen baj, és az mindig elérhető, legfeljebb megcseréljük a kimenetet)
 - OR-nél nincs ilyen baj, hiszen minden további nélkül mehet 1 fölé (a visszszámolt valószínűség mindig 0 és 1 között lesz)
- Az OR nem „összeönthető” (collapsible)
- A kettő közel van egymáshoz, ha a végpontok ritkák (mert ha $p \rightarrow 0$, akkor $\frac{p}{1-p} \rightarrow p$)

A végpontok jellegei: folytonos végpont

Folytonos végpont (pl. vérnyomás): eloszlásokat kapunk

- Jellemző mérőszám: az átlag változott-e, $MD = \bar{x}_C - \bar{x}_T$
- Érzéketlen arra, ha csak a szóródás változik (v.ö. TAR, TBR a diabetológiában)
- Ferde eloszlásoknál gyakran jobb a medián
 - Vigyázat, ez adatprezentációs célokra értendő, adatfeldolgozásra jó lehet az átlag az egyéb paraméterek (pl. mintanagyság) függvényében
- Ha más a skála, akkor használatos a standardizált MD (Cohen-féle d): $SMD = \frac{\bar{x}_C - \bar{x}_T}{s}$, mert mértékegységfüggetlen (szórás-egységekben méri a változást)
- Nagyon óvatosan a binarizálással (dichotomizálással)
 - Ismét csak: adatprezentálásra jó lehet (hány százaléknak van magas vérnyomása), de feldolgozáshoz a nyers adatot használjuk
 - Ha muszáj binarizálni (gyógyszert vagy adunk vagy nem – bár sokszor ez sem igaz, lehet kisebb vagy nagyobb dózist, ritkábban vagy gyakrabban stb.), akkor az elemzési folyamat lehető legkésőbbi pontján tegyük

A végpontok jellegei: eltelt idő jellegű végpont

Folytonos, de idő; eltelt idő jellegű végpont: túlélési görbéket kapunk

- Szóhasználat (függetlenül a konkrét helyzettől) egységesen: túlélési idő
- Túlélési görbe: a még életben lévők (végpontot el nem érők) aránya az eltelt idő függvényében
- (Elvileg – sokaságban – a valószínűség, ezt mintából aránnyal becsüljük)
- A probléma: cenzorálás
- „A beteg a 2 éves viziten még megjelent, de további információt nem tudtunk róla szerezni”
- Naiv logika: nem tudjuk, hogy mikor halt meg (érte el a végpontot), így *nyilván* nem tudjuk felhasználni ebben a vizsgálatban, akár csak a túlélési görbe meghatározásában, hiszen ahhoz kellene tudnunk, hogy mennyi idő telt el a halálig
- Csakhogy ez nem igaz!

- Az igaz, hogy nem tudjuk *biztosan*, hogy mi a végpontig eltelt idő
- De az nem igaz, hogy nem tudunk semmit!
- Valamit igenis tudunk: azt, hogy bármennyi is ez az idő, de *több mint 2 év*
- Ezt az információt fel kellene tudnunk használni
- Ezt hívják (jobb oldali) cenzorálásnak: részleges információnk van, csak azt tudjuk, hogy az érték nagyobb mint egy bizonyos szám
- Minimum hatásossági okokból fel kellene minden információt használni, ami rendelkezésre áll, de az ilyenek elhagyása torzítást is okozna
- Fontos: cenzorálás nem csak a fenti esetben van! Két alapeset létezik:
 - Loss to follow-up: mint fent, egyszerűen eltűnik az alany, nincs róla teljes utánkövetésünk
 - Teljes utánkövetésünk van az alanyról, csak épp végetért a kutatás mielőtt elérte volna a végpontot (ez is ugyanúgy cenzorálás!)

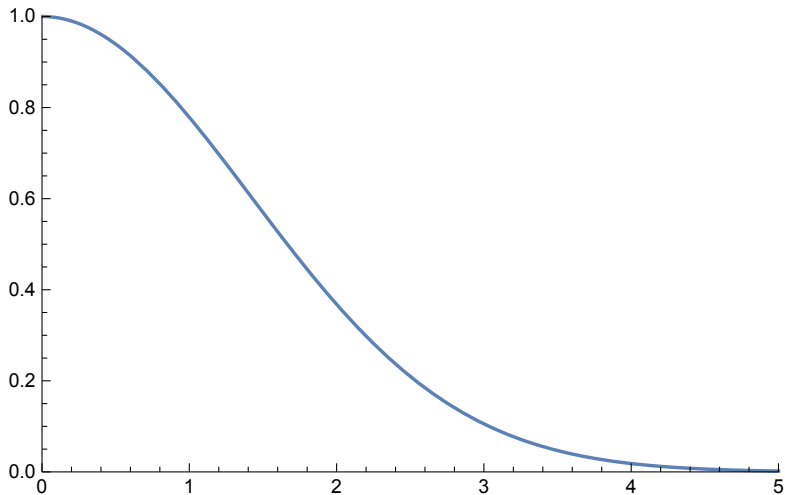
Túlélési görbe becslése cenzorálás jelenlétében

- Érdekes módon nem olyan nagyon régen, 1958-ban oldotta meg Edward Kaplan és Paul Meier
- A példa kedvéért: 100-ból az első évben meghalnak 5-en, a másodikban 2-en, a harmadikban 3-an; mennyi a túlélés a harmadik év végén?
- Kézenfekvő számolás: túléltek 90-en, összesen voltak 100-an, tehát $\frac{90}{100} = 0,9$, azaz 90%
- De számolhatunk másképp is: annak a valószínűsége, hogy valaki megéli a harmadik évet egyenlő annak a valószínűsége, hogy megéli az elsőt *szorozva* azzal, hogy megéli a másodikat *feltéve*, hogy megélte az elsőt *szorozva* azzal, hogy megéli a harmadikat *feltéve*, hogy megélte a másodikat
- Vagyis: $\frac{95}{100} \cdot \frac{93}{95} \cdot \frac{90}{93} = 0,9$ – tehát ugyanúgy 90%!
- Azaz: cenzorálás nélkül a két számítás ugyanarra jut
- De mi van, ha van cenzorálás, pl. a első évben kicenzorálódik 1, a másodikban 4?

Túlélési görbe becslése cenzorálás jelenlétében

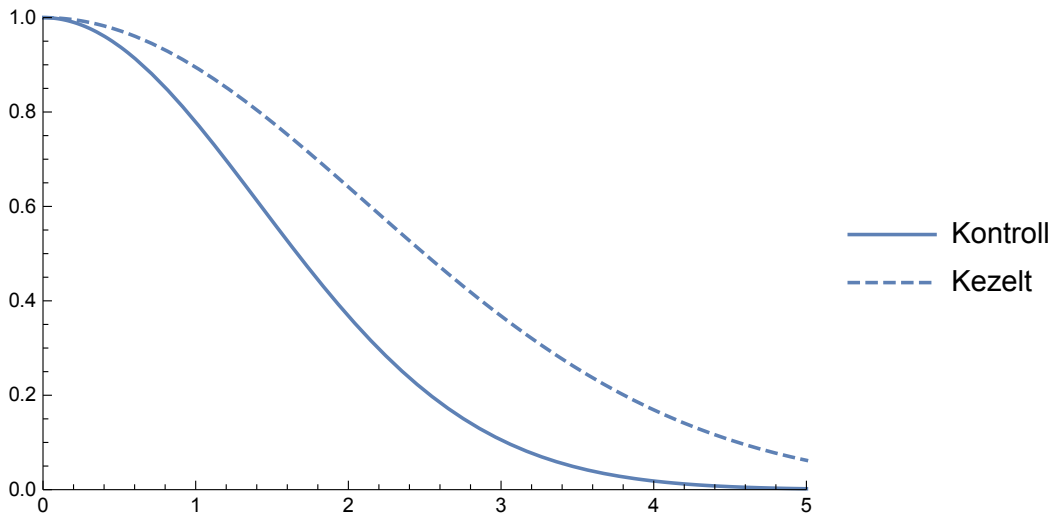
- Itt jön a kulcslépés: a második számításba beépíthető ez!
- Egyszerűen annyit kell tenni, hogy a halálozások számához nem nyúlunk, de alul amivel osztunk, azt *lecsökkentjük* a cenzorálással!
- Ugyanis osztani a „kockázatnak kitett populáció”-val kell – csak hogy ez a cenzorálás lévén is csökken (nem csak a halálozás révén), hiszen aki kicenzorálódott, az már nem tud meghalni!
- Tehát az többé nem a „nem meghaltak”, hanem a „nem meghaltak és nem cenzorálódottak” száma az, amivel osztunk
- A példában: $\frac{95}{100} \cdot \frac{92}{94} \cdot \frac{85}{88} = 0,898$, azaz 89,8%
- (Ezt persze le kell vezetni, pl. ML-elven is megtehető)
- Túlélési görbén: tipikus egy apró metsző vonal ott, ahol cenzorálás van
- Ettől még a görbe tovább fut, ugyanúgy vízszintesen mintha mi sem történt volna – csak épp a következő halálozás már „jobban fog számítani”!

Túlélési görbe



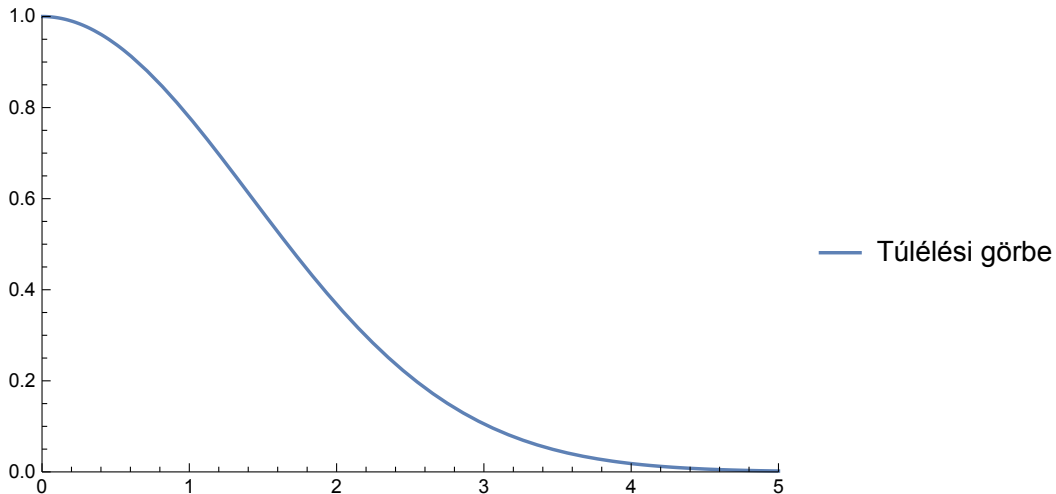
— Túlélési görbe

Túlélési görbék

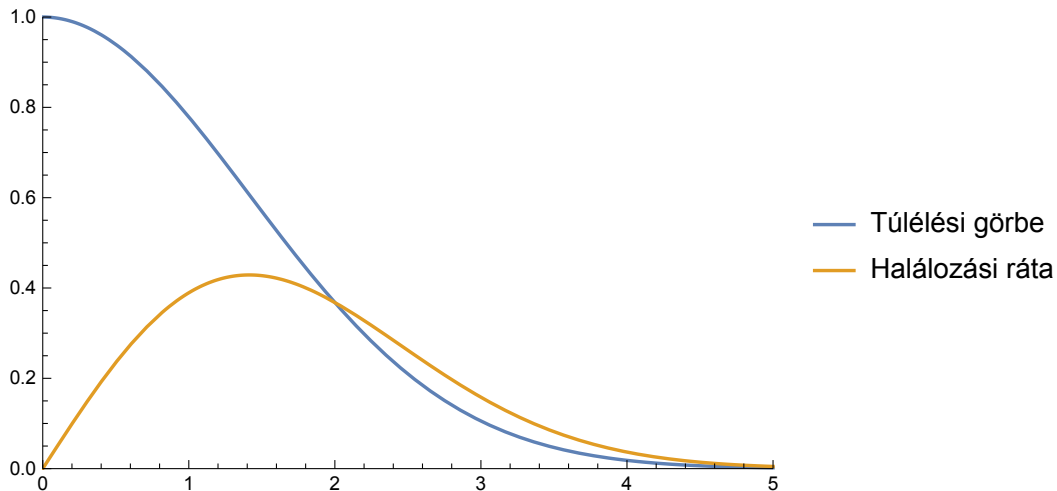


- Adott idejű, például 1 éves, túlélés (ezzel lényegében binarizáljuk a végpontot)
- Medián túlélés (szintén levágjuk, csak most „vízszintesen”)
- Vagy: hazard változott-e
- Alternatív mérőszámok (RMST, cure modellek)

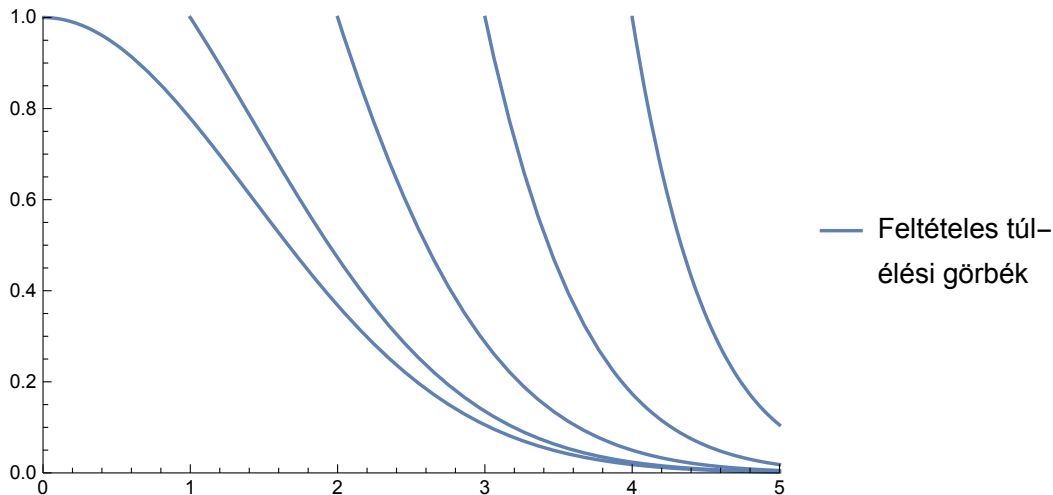
Eltelt idő jellegű végpont: a hazard fogalmának illusztrálása



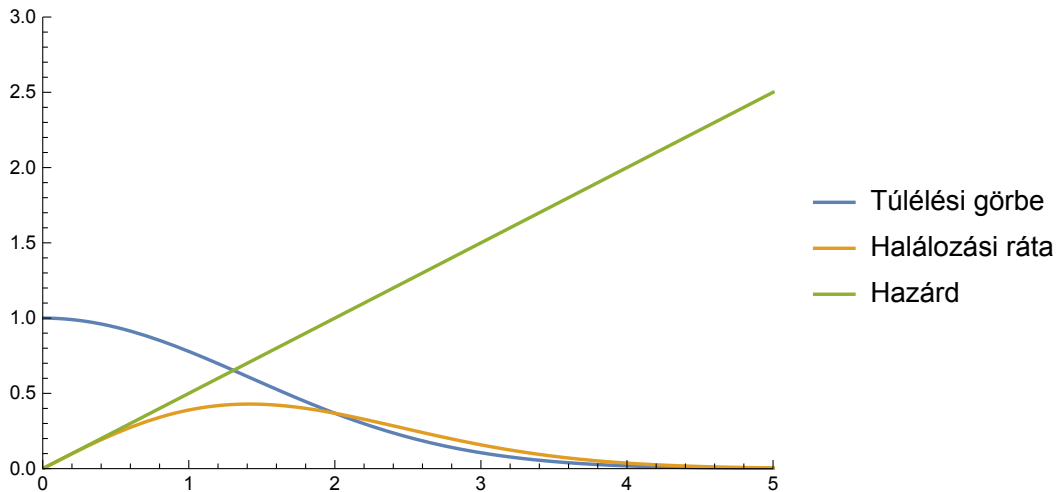
Eltelt idő jellegű végpont: a hazard fogalmának illusztrálása



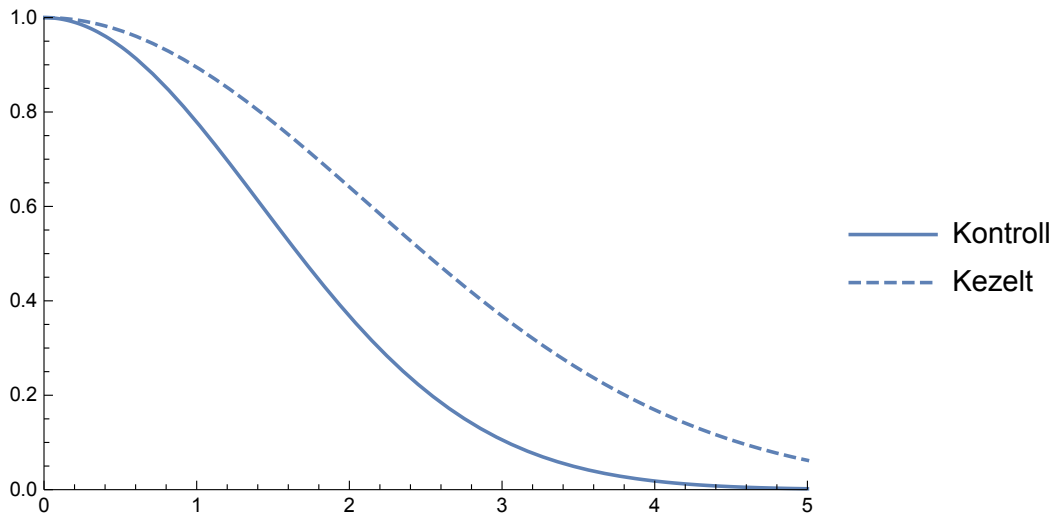
Eltelt idő jellegű végpont: a hazard fogalmának illusztrálása



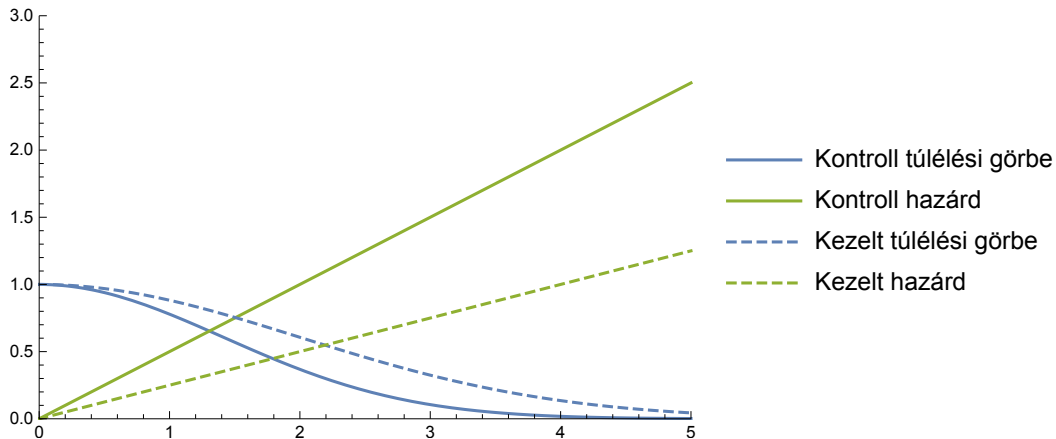
Eltelt idő jellegű végpont: és akkor most jön az összehasonlítás



Eltelt idő jellegű végpont: és akkor most jön az összehasonlítás



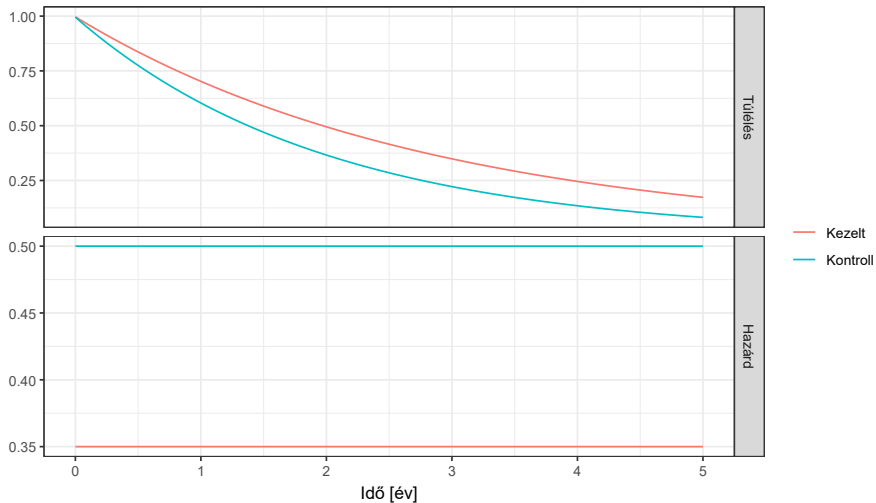
Eltelt idő jellegű végpont: és akkor most jön az összehasonlítás



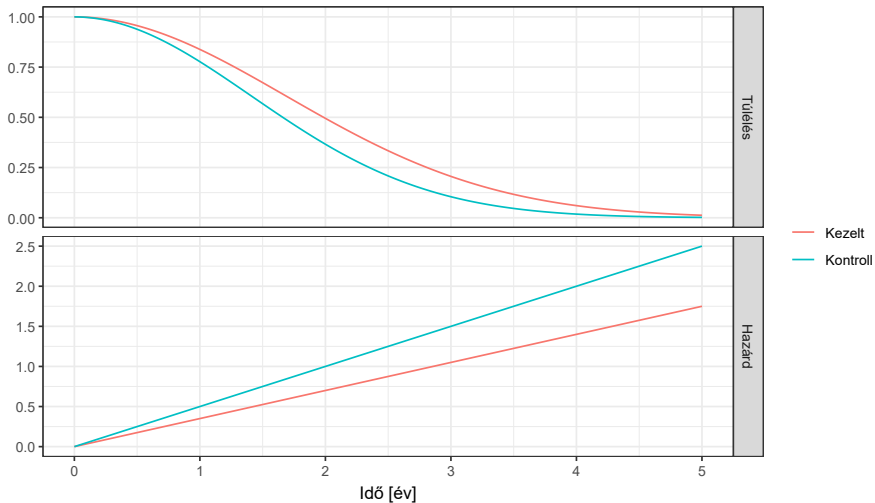
A HR definíciója

- Nagyon fontos: nem kell, hogy a házárdfüggvény ilyen egyenes legyen (ez csak véletlenül volt így az előző ábrákon)...
- ...de erre nincs is szükség!
- Csak annyi a fontos, hogy a kezelt csoport házárdja *minden időpillanatban* ugyanannyiszorosa legyen a kontrollcsoportének (a kettő hányadosa minden időpillanatban ugyanannyi legyen)
- Házárdhányados (hazard ratio, HR): Hányszorosa a kezelt csoport házárdja a kontrollcsoporténak?
- Na de melyik időpontban?
- Mindegy! (Mert mindegyik időpontban ugyanannyiszorosa)
- Ezt hívjuk úgy, hogy a házárdok proporcionálisak, a HR csak ilyen esetben értelmezhető egyértelműen

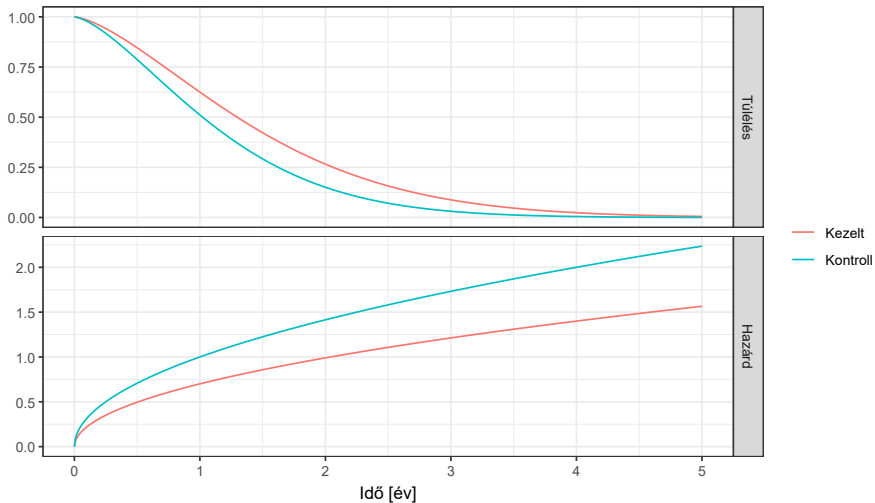
Pár illusztráció proporcionális hazárdokra: $HR = 0,7$



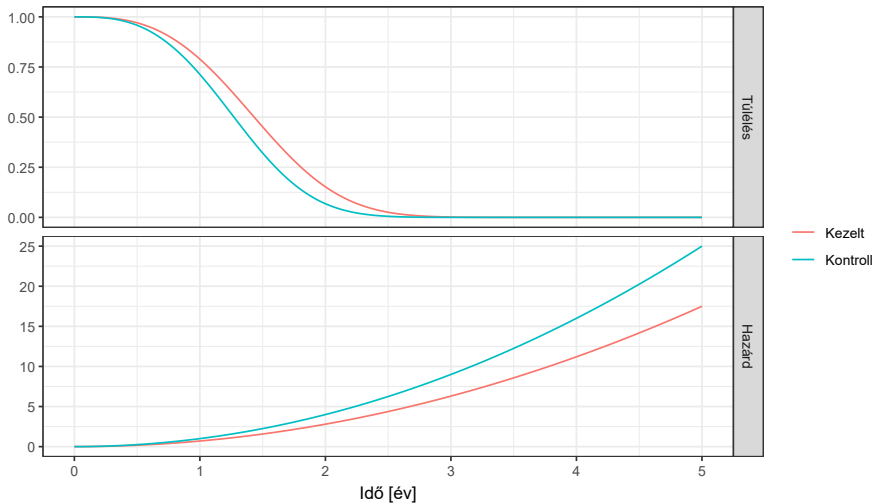
Pár illusztráció proporcionális hazárdokra: $HR = 0,7$



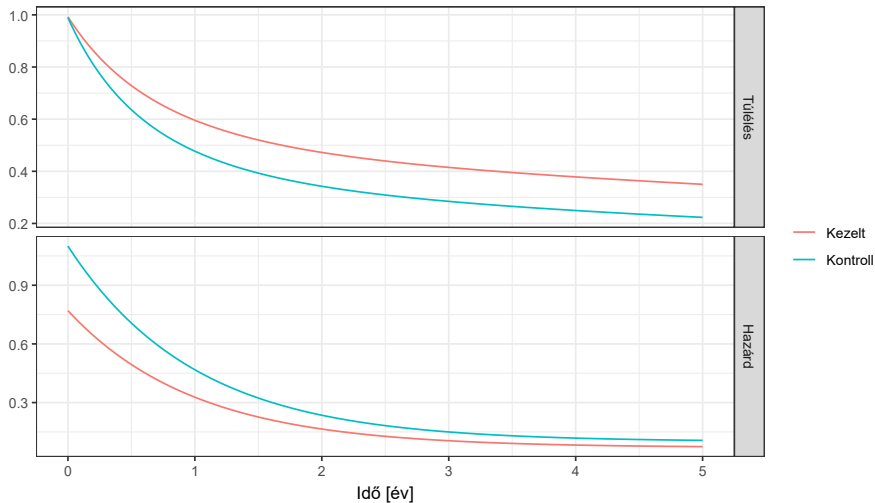
Pár illusztráció proporcionális hazárdokra: $HR = 0,7$



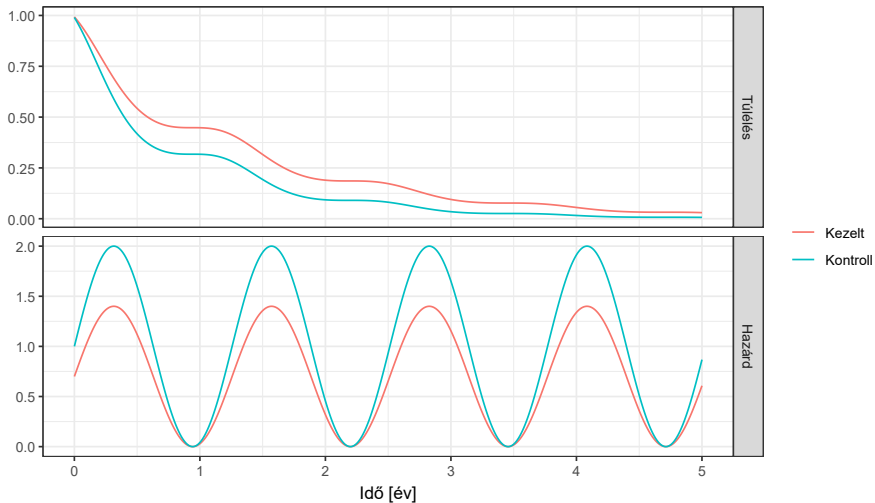
Pár illusztráció proporcionális hazárdokra: $HR = 0,7$



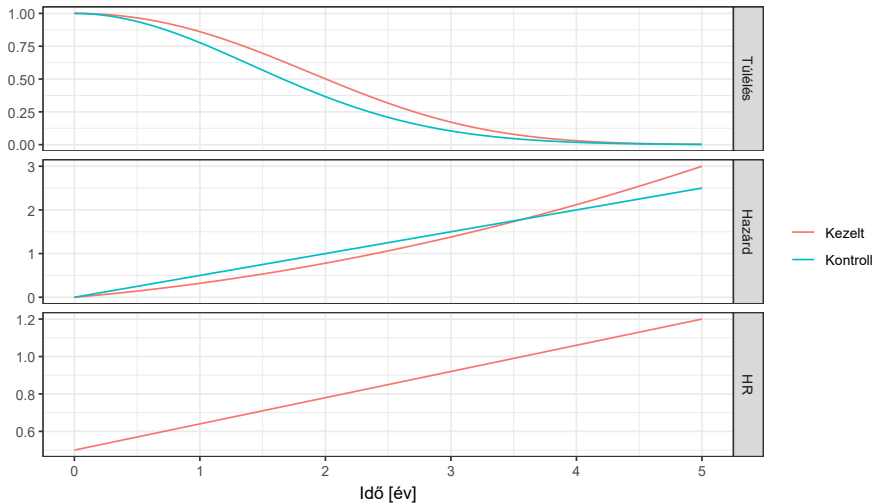
Pár illusztráció proporcionális hazárdokra: $HR = 0,7$



Pár illusztráció proporcionális hazárdokra: $HR = 0,7$



De: nem-proporcionális hazardok



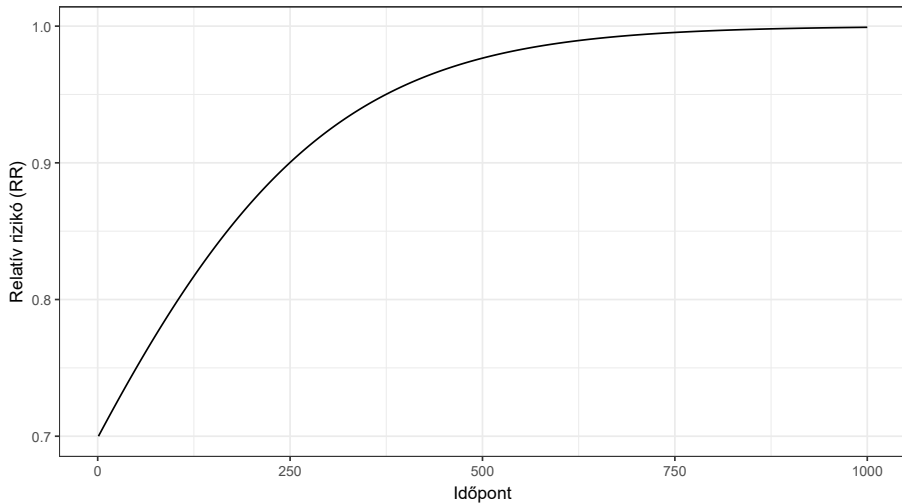
- $HR = 0,7$
 - A kezelés 30%-kal csökkenti a halálozás kockázatát?
 - A kezelés 30%-kal csökkenti a halálozás kockázatát adott időtávon belül?
- A probléma a kockázat szó: a hazard, ha nagyon akarjuk, valóban kockázat, de egy *rendkívül* speciális értelemben (ami nem egyezik azzal, amit hétköznapiilag értünk „kockázat” alatt!)
- Pillanatnyi vs. időintervallumra vonatkozó kockázat

A pillanatnyi és a szokásos kockázat

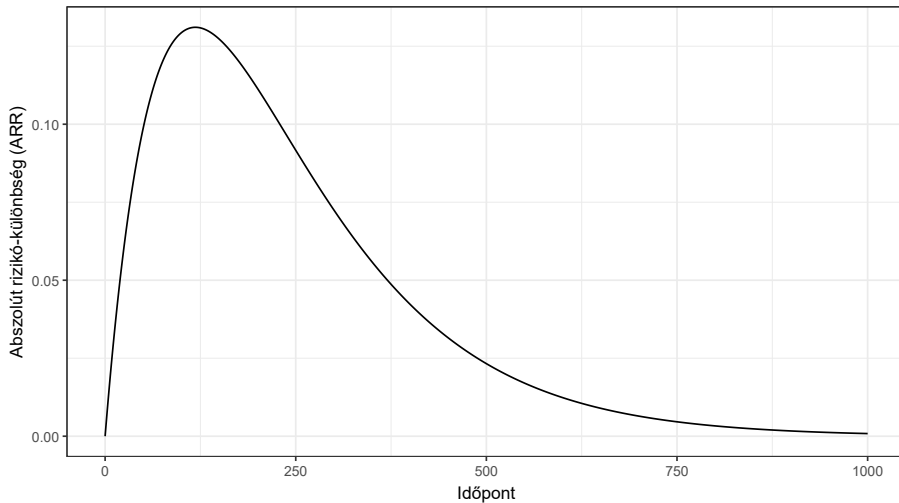
Baseline hazard 1% / nap (azaz exponenciális túlélés), $HR = 0,7$ proporcionális hazardddal

Időpont	Kontroll	Kezelt	Relatív rizikó
0	1000	1000	
1	990	993	0,700
2	980	986	0,701
3	970	979	0,702
4	961	972	0,703
5	951	965	0,704
6	941	959	0,705
7	932	952	0,706
⋮	⋮	⋮	⋮
363	26	78	0,947
364	26	78	0,947
365	26	77	0,947

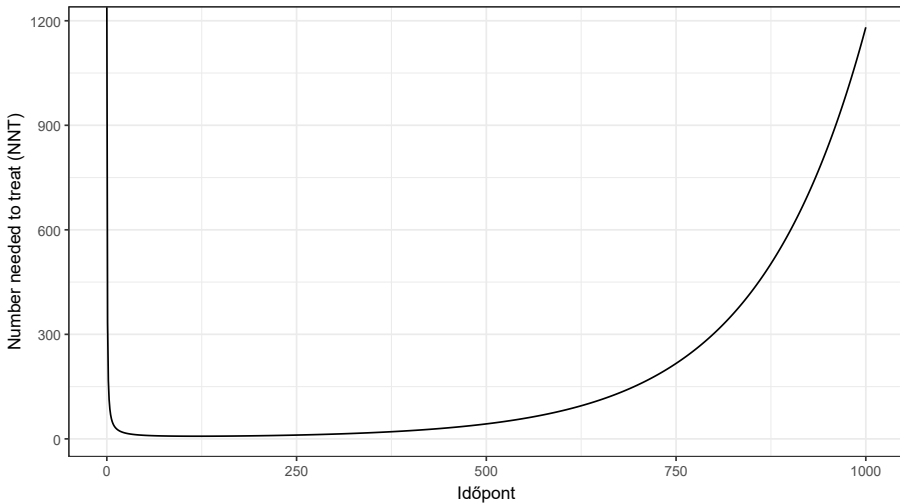
A pillanatnyi és a szokásos kockázat: RR



A pillanatnyi és a szokásos kockázat: ARR



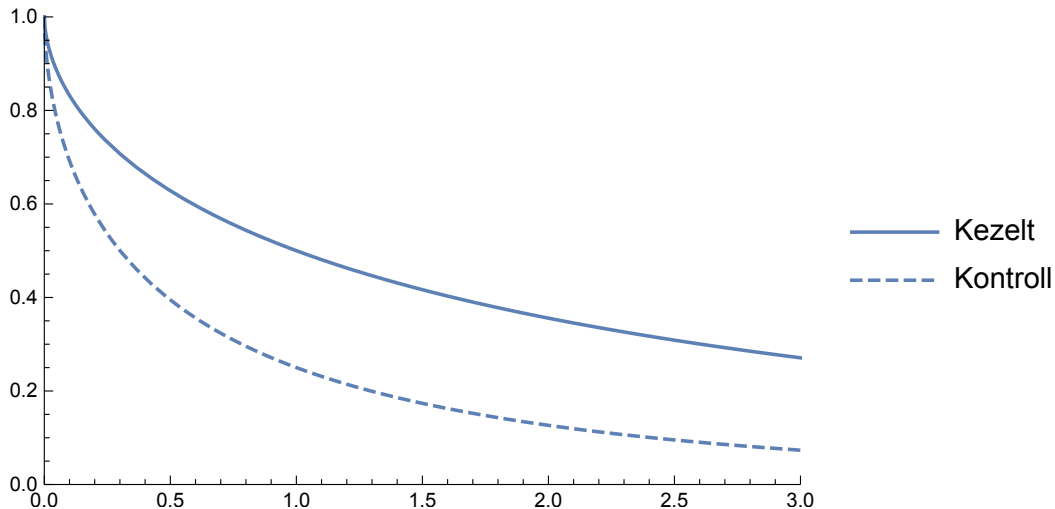
A pillanatnyi és a szokásos kockázat: NNT



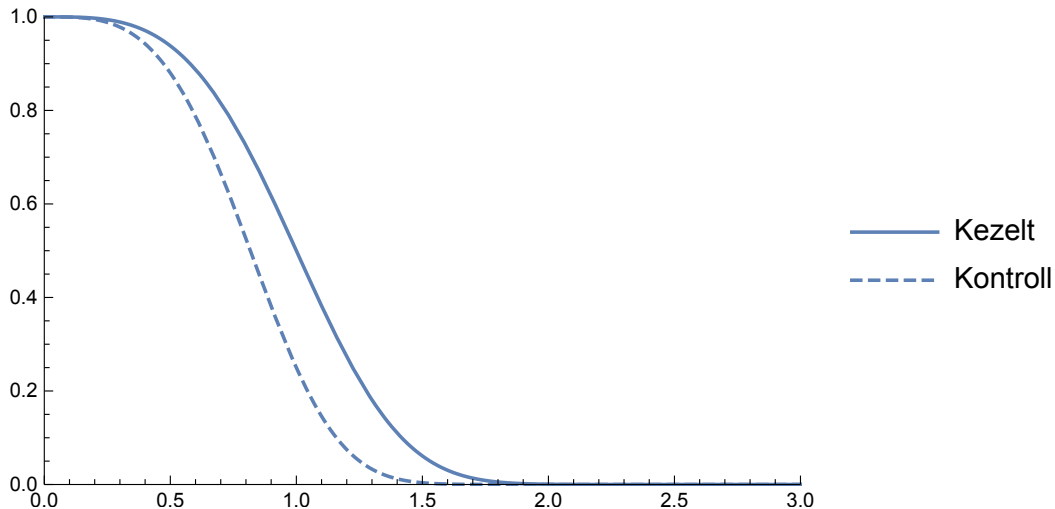
- Sokszor valamiféle „sebességként” kezelik (első ránézésre érthető módon, mi is azt mondtuk, hogy „milyen gyorsan” fogynak a betegek), például ilyeneket mondva:
 - kétszer hamarabb vagy kétszer gyorsabban halnak meg kezelés nélkül a betegek...
 - kétszer annyian halnak meg adott időpontban...
 - kétszer annyian halnak meg adott időpontig...
 - kétszer valószínűbb, hogy meghalnak kezelés nélkül...
- Ezen megfogalmazások mindegyike hibás!
- A sebesség-analógia félrevezető: a sebesség értelmezhető egy konkrét autóra, de a „milyen gyorsan fogynak” megközelítés csak betegek egy csoportjára: az teljesen értelmezhető, hogy a betegek ezen csoportja ebben az időpontban 1%/nap gyorsasággal fogy, azaz hal meg (mert például 1 nap alatt 100-ból 1 halt meg), de ez egy konkrét betegre értelmetlen, hiszen ő nem tud 1 nap alatt 1%-ig meghalni...

- A HR a túlélési görbe alakjától is függ
- Végeredményben a klinikailag releváns hatást egyáltalán nem biztos, hogy úgy jellemzi, ahogy azt intuitíve várnánk

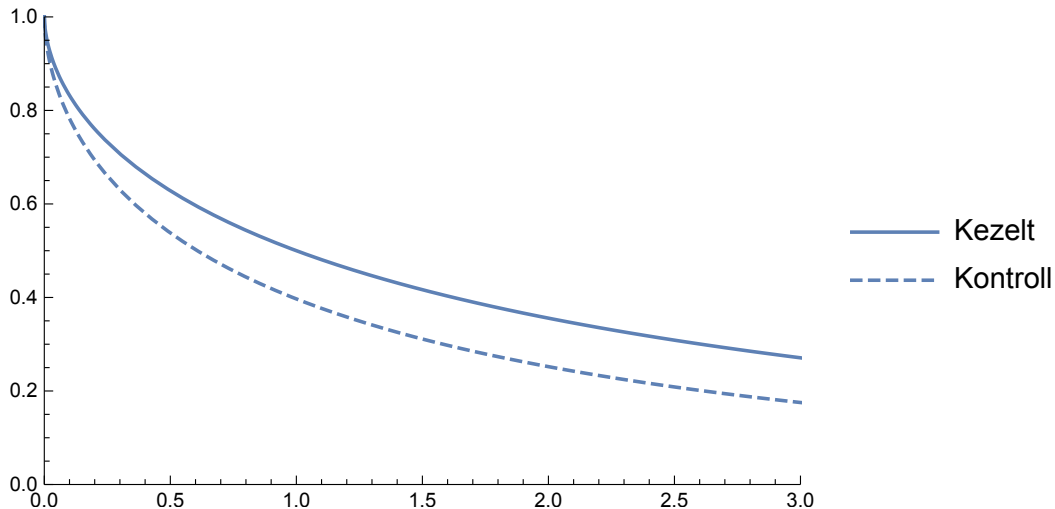
Remek kezelés: $HR=0,5$



Ugyanolyan (???) remek kezelés: $HR=0,5$



Gyengécske kezelés: $HR=0,75$



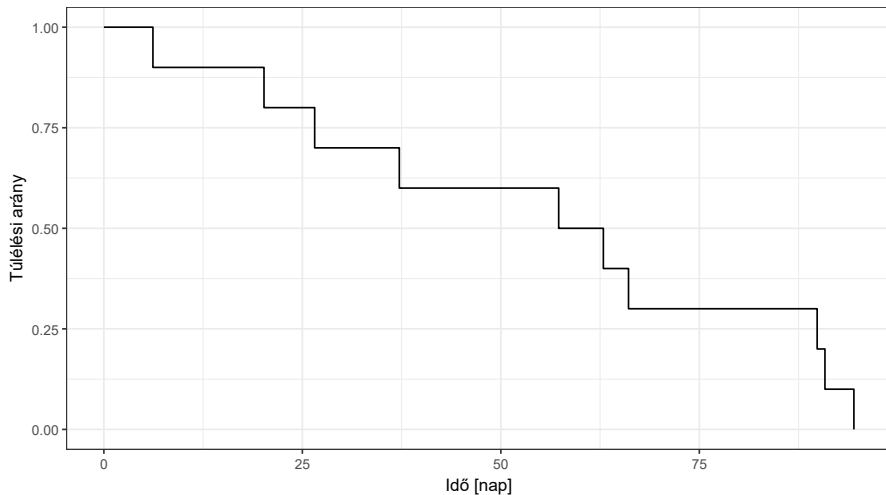
- Felhasználja a teljes túlélési görbét (nem csak egy pontját, mint a fix idejű túlélési arány vagy a medián túlélési idő)
- De kérdés, hogy a proporcionalitás teljesül-e
- Figyelni kell az interpretációval, ahogy az előbb láttuk
- Függ a túlélési görbe alakjától is
- A HR relatív mutató – problémás, hogy mennyire jellemzi jól a hatánagyságot („klinikai előny”)
- (A fix idejű túlélési arány és a medián túlélési idő abszolút mutató!)

Céljaink lehetnek:

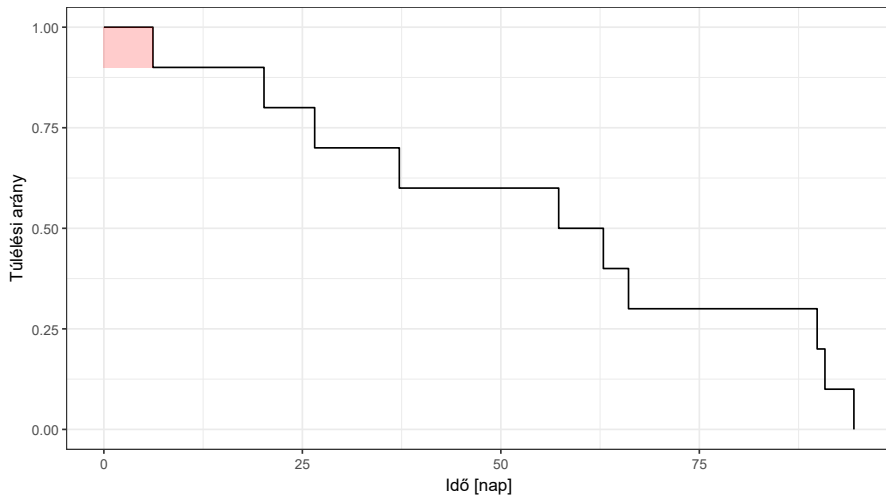
- Feloldani a proporcionalitás szükségességét
- Interpretálható (abszolút jellegű) mutatót adni, mely a klinikai hasznot jobban jellemzi, jobban lefordítható kézzelfogható klinikai előnyre

- Miért nem nézzük az átlagos túlélési időt?
- Először értsük ezt meg: mi az az átlagos túlélési idő? Hol jelenik ez meg?
- Kezdjük az átlag definíciójával: össze kell adni minden betegre, hogy meddig élt, és ezt az összeget el kell osztani a betegek számával ($\frac{t_1+t_2+\dots+t_{10}}{10}$)
- Avagy minden beteg túlélési idejét el kell osztani a betegek számával és ezeket összeadni ($\frac{t_1}{10} + \frac{t_2}{10} + \dots + \frac{t_{10}}{10}$)
- Avagy minden beteg túlélési idejét meg kell szorozni a betegek számának reciprokéval és ezeket összeadni ($\frac{1}{10}t_1 + \frac{1}{10}t_2 + \dots + \frac{1}{10}t_{10}$)
- Nézzünk egy egyszerű példát: 10 beteg (minden beteg túlélési idejét 0,1-gyel kell szorozni), nincs cenzorálás, 0-ba megy a túlélési görbe

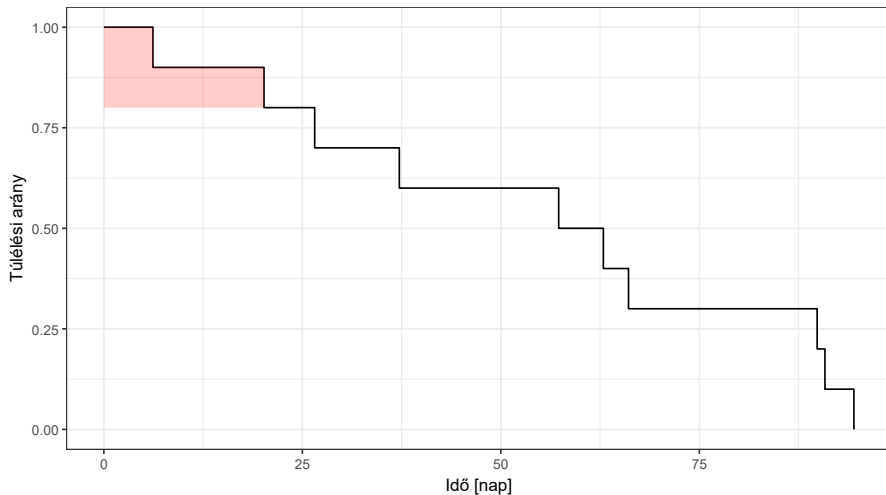
Átlagos túlélési idő szemléltetése



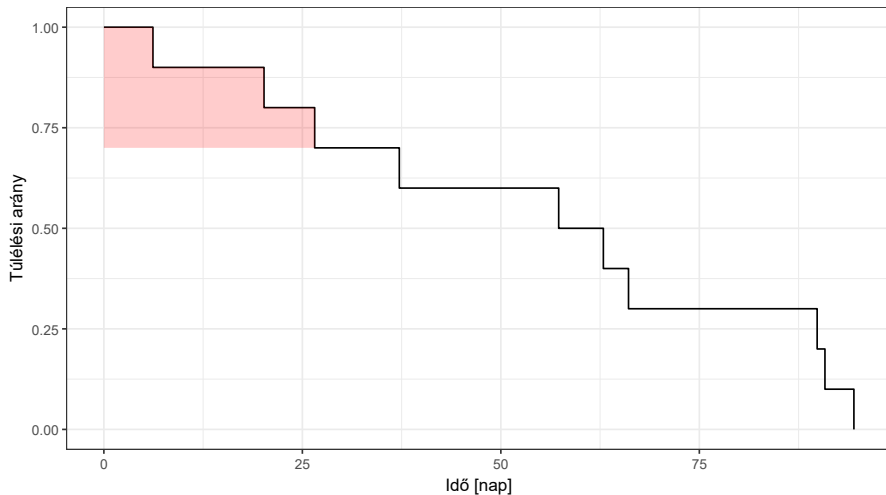
Átlagos túlélési idő szemléltetése



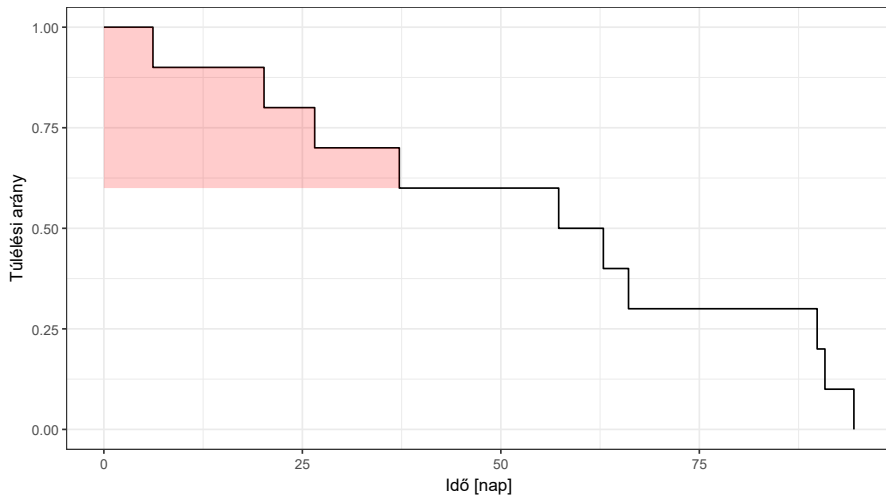
Átlagos túlélési idő szemléltetése



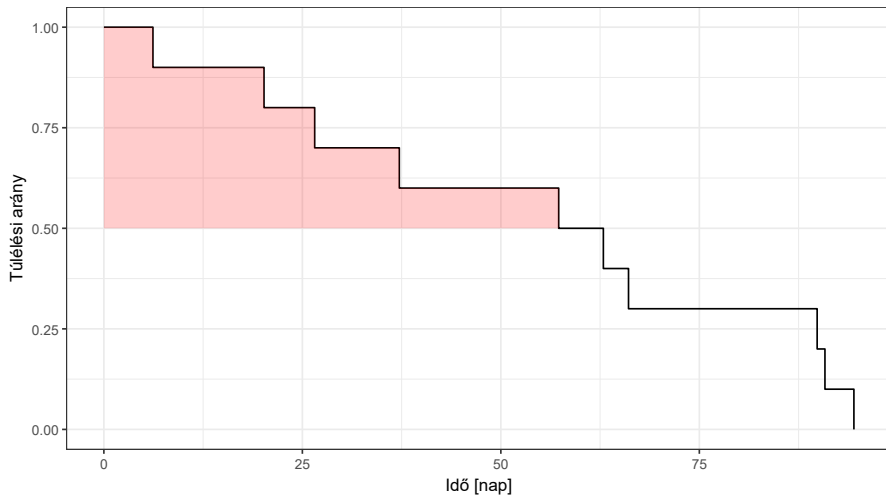
Átlagos túlélési idő szemléltetése



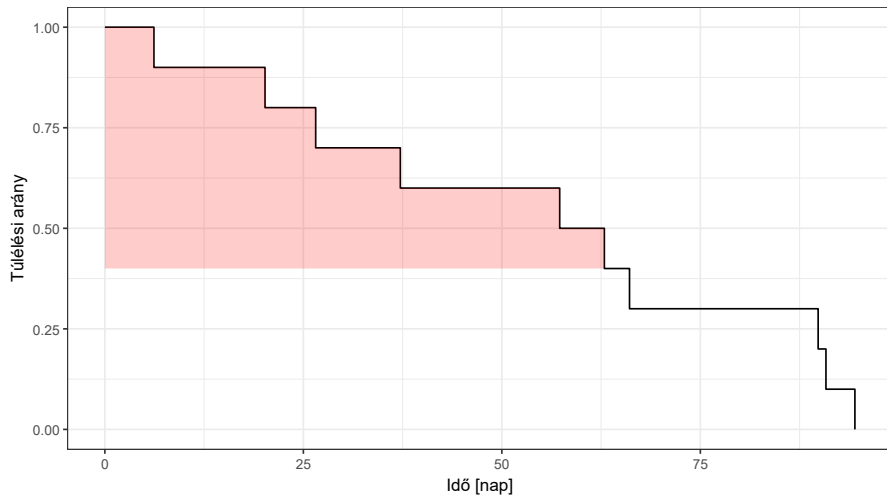
Átlagos túlélési idő szemléltetése



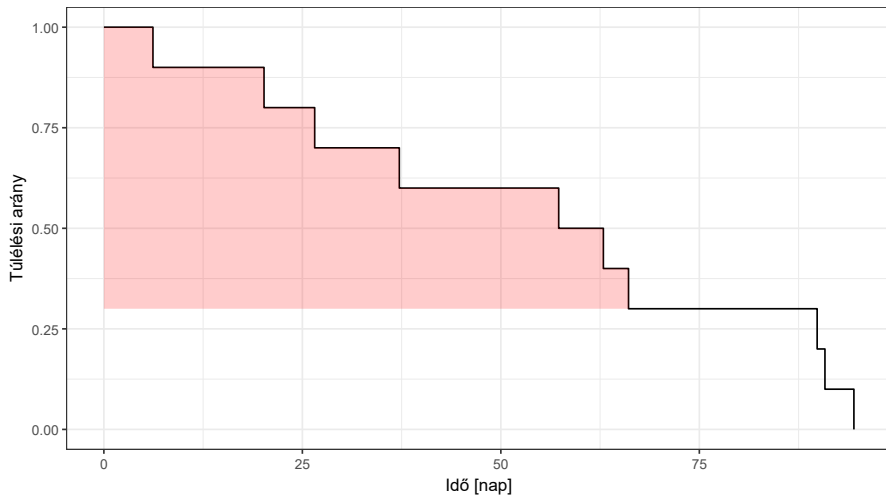
Átlagos túlélési idő szemléltetése



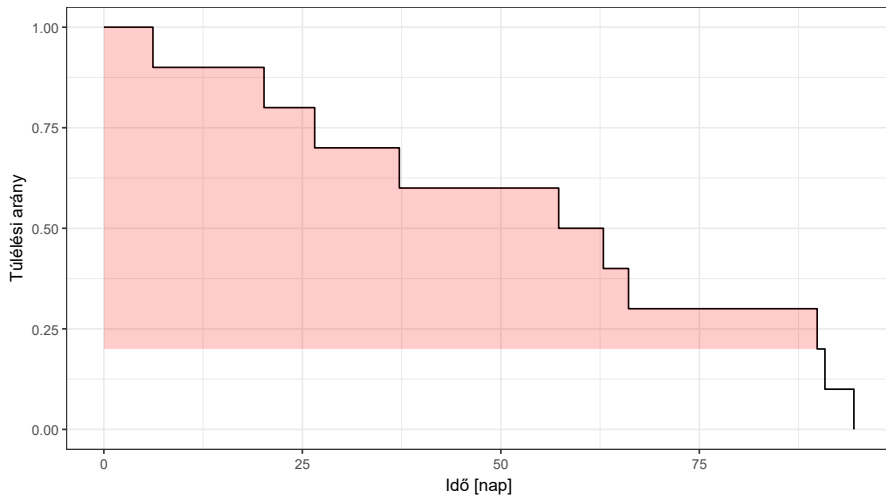
Átlagos túlélési idő szemléltetése



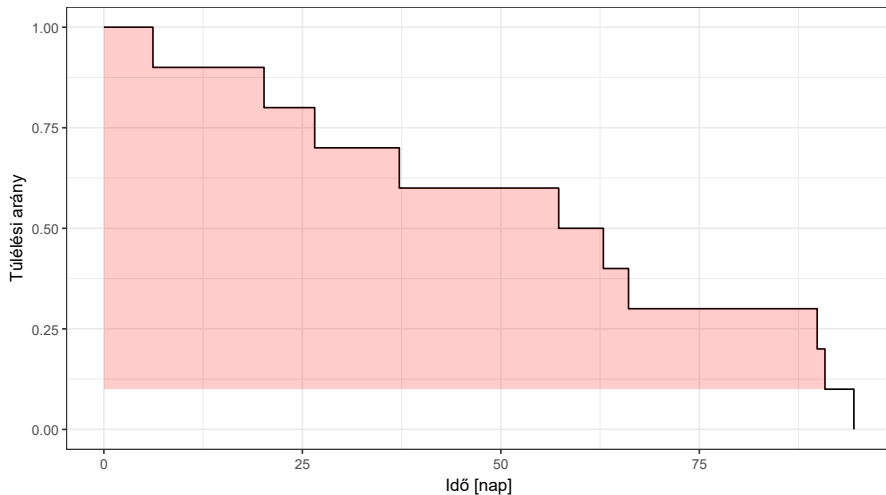
Átlagos túlélési idő szemléltetése



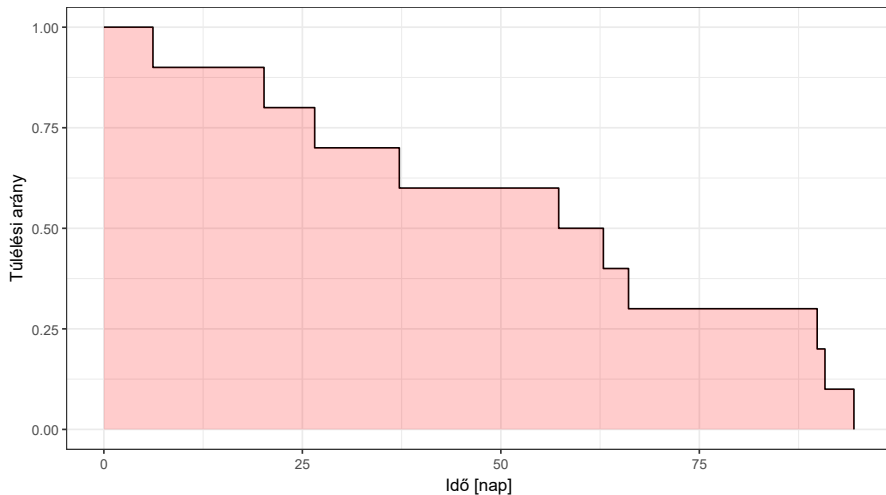
Átlagos túlélési idő szemléltetése



Átlagos túlélési idő szemléltetése



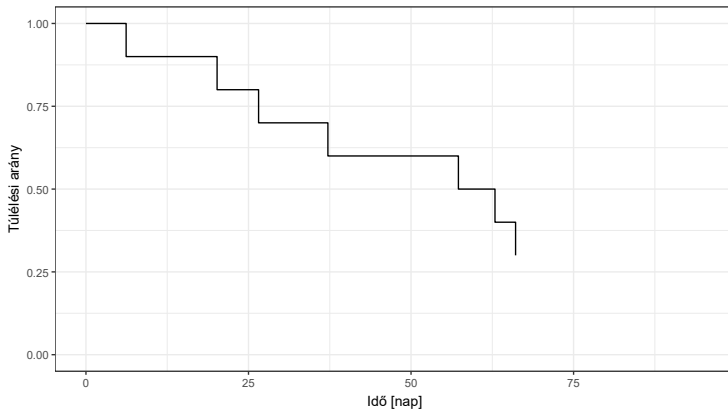
Átlagos túlélési idő szemléltetése



- ...azaz az átlagos túlélési idő épp a görbe alatti terület!
- (A görbe alatti területet ugyan úgy szokták szemléltetni, hogy függőleges oszlopokat húznak egymás mellé, de a végeredmény szempontjából mindegy)
- Nem kevésbé érdekes a különbség az átlagos túlélési időben: a két túlélési görbe közti terület!
- „Hány hónapot adunk átlagosan a betegeknek ezzel a kezeléssel?”
- (100 beteg összesen 500 hónappal tovább élt, akkor teljes joggal mondhatjuk, hogy egy beteg *átlagosan* 5 hónapot nyert)
- Gyakran szokták kérdezni, mert kézenfekvő (klinikai hasznót jól mutató metrika!), de meg szoktak lepődni, hogy ez az átlagos túlélési időre vezet vissza
- A hányados is releváns relatív mutatóként

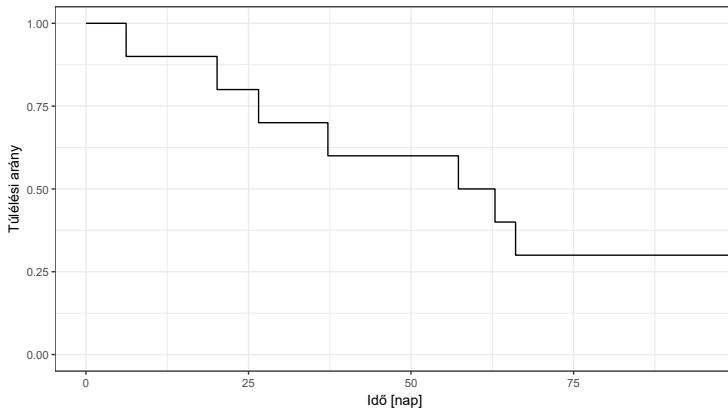
A nem teljes utámkövetés problémája

Mi van, ha a túlési görbe nem megy 0-ba (mert cenzorálásokkal ér véget)? Úgy szokták ábrázolni, mintha ott „véget érne” a túlélési görbe...



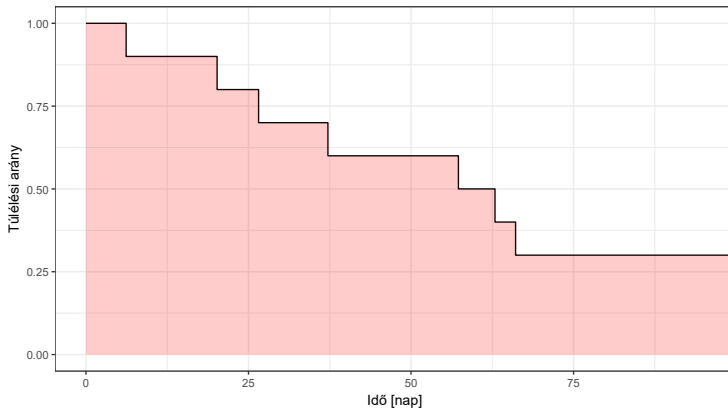
A nem teljes utámkövetés problémája

...valójában persze a görbe nem tud véget érni – az megy tovább, mégpedig vízszintesen!



A nem teljes utámkövetés problémája

Emiatt a görbe alatti terület végtelen lesz!

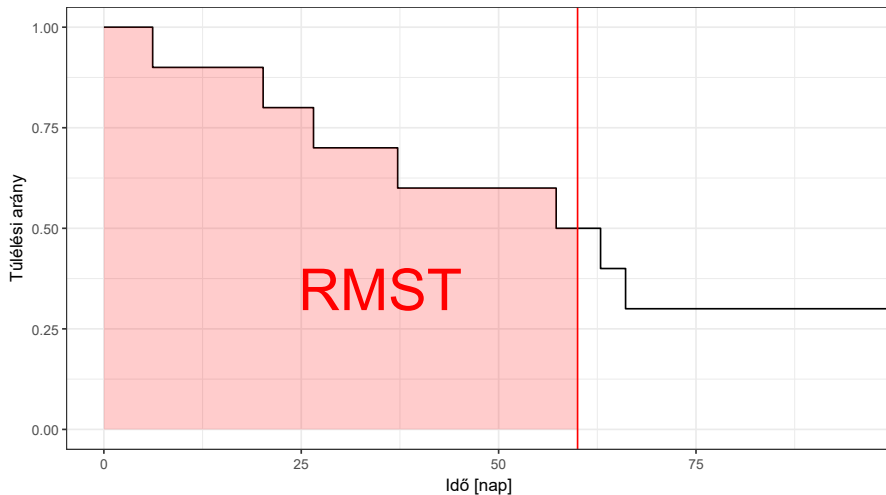


Az átlagos túlélési idő problémája

- A probléma: ha a túlélési görbe nem megy nullába, akkor az átlagos túlélési idő végtelen!
- Jobbra „kifut végtelenbe” a sáv (a még élő betegek olyan, mintha végtelen túlélési idővel bírnának)
- Pont ezért jó mellel a medián: ha a számok 49%-a végtelen, a medián akkor is értelmezett, az átlag már akkor sem, ha csak egyetlen egy végtelen!
- Ötlet: csak egy adott időpontig bezárólag adjuk össze!
- Restricted Mean Survival Time (RMST) fogalma
- Az előzmények régebbiek, de az orvosi irodalomban a 2010-es évek elején jelent igazán meg (Royston és Parmar)

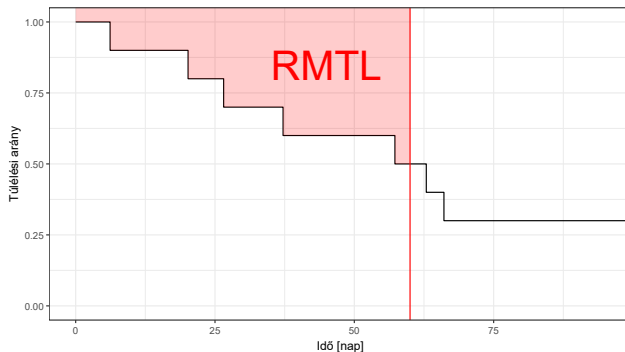
Royston P, Parmar MKB: The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt. Stat Med. 2011, 30: 2409-2421. 10.1002/sim.4274. Royston P, Parmar MK. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. BMC Med Res Methodol. 2013;13:152.

Az RMST szemléltetése



Az RMTL fogalma

Az átlagos elvesztett idő (RMTL), vagy – két görbe esetén – azok hányadosa szintén releváns lehet:



Hajime Uno, Brian Claggett, Lu Tian et al. Moving Beyond the Hazard Ratio in Quantifying the Between-Group Difference in Survival Analysis. J Clin Oncol. 2014 Aug 1; 32(22): 2380–2385. doi: 10.1200/JCO.2014.55.2208.

- Nem függ olyan feltevéstől mint a kockázatok arányossága, működik a nélkül is („modell-mentes”)
- Ráadásul ha fennáll, akkor a szignifikanciák nagyon hasonlítanak a kockázatokon alapuló módszerekéhez

TABLE 1 HR, RMSTs, and RMTL Ratios for All-Cause Mortality in Pharmaceutical Trials									
Trial (Ref. #)	Treatment	HR Reported	HR p Value	T* (Months)	RMST (95% CI) (Months)	RMTL (95% CI)	RMTL p Value	ΔRMST (95% CI)	ΔRMST p Value
CONSENSUS (24)	Enalapril	0.73†	0.003	12	8.7 (7.9 to 9.5)	0.6 (0.4 to 0.8)	0.001	2.2 (1.0 to 3.4)	<0.001
	Placebo	Ref.		12	6.5 (5.6 to 7.4)	Ref.		Ref.	
SOLVD TREATMENT (35)	Enalapril	0.84‡ (0.74 to 0.95)	0.004	48	38.1 (37.2 to 38.9)	0.8 (0.7 to 0.9)	0.003	1.9 (0.7 to 3.2)	0.002
	Placebo	Ref.		48	36.2 (35.3 to 37.1)	Ref.		Ref.	
RALES (25)	Spironolactone	0.70 (0.60 to 0.82)	0.001	34	26.2 (25.4 to 27.0)	0.8 (0.7 to 0.9)	<0.001	2.2 (1.1 to 3.4)	<0.001
	Placebo	Ref.		34	23.9 (23.1 to 24.8)	Ref.		Ref.	
CIBIS-II (26)	Bisoprolol	0.66 (0.54 to 0.81)	<0.001	24	22.0 (21.7 to 22.3)	0.7 (0.6 to 0.8)	<0.001	0.9 (0.4 to 1.4)	<0.001
	Placebo	Ref.		24	21.1 (20.8 to 21.5)	Ref.		Ref.	
MERIT-HF (36)	Metoprolol	0.66 (0.53 to 0.81)	<0.001	18	17.1 (17.0 to 17.4)	0.7 (0.6 to 0.8)	<0.001	0.4 (0.2 to 0.7)	<0.001
	Placebo	Ref.		18	16.7 (16.5 to 16.8)	Ref.		Ref.	
SHIFT (37)	Ivabradine	0.90 (0.80 to 1.02)	0.092	30	27.3 (27.0 to 27.5)	0.9 (0.8 to 1.0)	0.052	0.3 (−0.01 to 0.70)	0.052
	Placebo	Ref.		30	26.9 (26.7 to 27.2)	Ref.		Ref.	
EMPHASIS-HF (38)	Eplerenone	0.76 (0.62 to 0.93)	0.008	36	32.7 (32.2 to 33.2)	0.8 (0.6 to 0.9)	0.010	1.0 (0.2 to 1.7)	0.009
	Placebo	Ref.		36	31.7 (31.1 to 32.2)	Ref.		Ref.	
PARADIGM-HF (39)	Sacubitril/valsartan	0.84 (0.76 to 0.93)	<0.001	41	36.2 (35.9 to 36.5)	0.9 (0.8 to 0.9)	<0.001	0.8 (0.3 to 1.3)	0.001
	Enalapril	Ref.		41	35.4 (35.1 to 35.8)	Ref.		Ref.	
DAPA-HF (40)	Dapagliflozin	0.83 (0.71 to 0.97)	NA*	24	22.3 (22.1 to 22.5)	0.8 (0.7 to 0.98)	0.028	0.3 (0.04 to 0.60)	0.028
	Placebo	Ref.		24	22.0 (21.7 to 22.2)	Ref.		Ref.	
VICTORIA (41)	Vericiguat	0.95 (0.84 to 1.07)	0.38	32	26.2 (25.7 to 26.6)	0.95 (0.9 to 1.1)	0.368	0.3 (−0.3 to 0.9)	0.367
	Placebo	Ref.		32	25.9 (25.4 to 26.3)	Ref.		Ref.	

*Calculated only for outcomes that were included in the trial's hierarchical testing strategy. †Based on risk reduction, estimated using life table method, ‡Based on relative risk, estimated using life table method.

Carlotta Perego, Marco Sbolli Claudia Specchia et al. Utility of Restricted Mean Survival Time Analysis for Heart Failure Clinical Trial Evaluation and Interpretation. JACC: Heart Failure. Volume 8, Issue 12, December 2020, Pages 973-983. doi: 10.1016/j.jchf.2020.07.005

- Nem függ olyan feltevéstől mint a kockázatok proporcionalitása, működik a nélkül is („modell-mentes”)
- Ráadásul ha fennáll, akkor a szignifikanciák nagyon hasonlítanak a kockázatokon alapuló módszerekéhez
- De közben a számérték jobban értelmezhető („várható túlélés hossza adott időpontig bezárólag”), klinikai haszonnak is (abszolút jellegű mutató!)
- Ebből fakadóan jobban összehasonlítható: értelmezhető a különbség (görbék közötti terület!), szintén abszolút mutató, és relatívvá is jól tehető, mert a kontrollcsoport RMST-je jó viszonyítási alap, így a relatív mutató – szemben a HR-rel – is jobban értelmezhető klinikai előnyként

TABLE 1 HR, RMSTs, and RMTL Ratios for All-Cause Mortality in Pharmaceutical Trials									
Trial (Ref. #)	Treatment	HR Reported	HR p Value	T* (Months)	RMST (95% CI) (Months)	RMTL (95% CI)	RMTL p Value	ΔRMST (95% CI)	ΔRMST p Value
CONSENSUS (24)	Enalapril	0.73†	0.003	12	8.7 (7.9 to 9.5)	0.6 (0.4 to 0.8)	0.001	2.2 (1.0 to 3.4)	<0.001
	Placebo	Ref.		12	6.5 (5.6 to 7.4)	Ref.		Ref.	
SOLVD TREATMENT (35)	Enalapril	0.84‡ (0.74 to 0.95)	0.004	48	38.1 (37.2 to 38.9)	0.8 (0.7 to 0.9)	0.003	1.9 (0.7 to 3.2)	0.002
	Placebo	Ref.		48	36.2 (35.3 to 37.1)	Ref.		Ref.	
RALES (25)	Spironolactone	0.70 (0.60 to 0.82)	0.001	34	26.2 (25.4 to 27.0)	0.8 (0.7 to 0.9)	<0.001	2.2 (1.1 to 3.4)	<0.001
	Placebo	Ref.		34	23.9 (23.1 to 24.8)	Ref.		Ref.	
CIBIS-II (26)	Bisoprolol	0.66 (0.54 to 0.81)	<0.001	24	22.0 (21.7 to 22.3)	0.7 (0.6 to 0.8)	<0.001	0.9 (0.4 to 1.4)	<0.001
	Placebo	Ref.		24	21.1 (20.8 to 21.5)	Ref.		Ref.	
MERIT-HF (36)	Metoprolol	0.66 (0.53 to 0.81)	<0.001	18	17.1 (17.0 to 17.4)	0.7 (0.6 to 0.8)	<0.001	0.4 (0.2 to 0.7)	<0.001
	Placebo	Ref.		18	16.7 (16.5 to 16.8)	Ref.		Ref.	
SHIFT (37)	Ivabradine	0.90 (0.80 to 1.02)	0.092	30	27.3 (27.0 to 27.5)	0.9 (0.8 to 1.0)	0.052	0.3 (−0.01 to 0.70)	0.052
	Placebo	Ref.		30	26.9 (26.7 to 27.2)	Ref.		Ref.	
EMPHASIS-HF (38)	Eplerenone	0.76 (0.62 to 0.93)	0.008	36	32.7 (32.2 to 33.2)	0.8 (0.6 to 0.9)	0.010	1.0 (0.2 to 1.7)	0.009
	Placebo	Ref.		36	31.7 (31.1 to 32.2)	Ref.		Ref.	
PARADIGM-HF (39)	Sacubitril/valsartan	0.84 (0.76 to 0.93)	<0.001	41	36.2 (35.9 to 36.5)	0.9 (0.8 to 0.9)	<0.001	0.8 (0.3 to 1.3)	0.001
	Enalapril	Ref.		41	35.4 (35.1 to 35.8)	Ref.		Ref.	
DAPA-HF (40)	Dapagliflozin	0.83 (0.71 to 0.97)	NA*	24	22.3 (22.1 to 22.5)	0.8 (0.7 to 0.98)	0.028	0.3 (0.04 to 0.60)	0.028
	Placebo	Ref.		24	22.0 (21.7 to 22.2)	Ref.		Ref.	
VICTORIA (41)	Vericiguat	0.95 (0.84 to 1.07)	0.38	32	26.2 (25.7 to 26.6)	0.95 (0.9 to 1.1)	0.368	0.3 (−0.3 to 0.9)	0.367
	Placebo	Ref.		32	25.9 (25.4 to 26.3)	Ref.		Ref.	

*Calculated only for outcomes that were included in the trial's hierarchical testing strategy. †Based on risk reduction, estimated using life table method, ‡Based on relative risk, estimated using life table method.

Carlotta Perego, Marco Sbolli Claudia Specchia et al. Utility of Restricted Mean Survival Time Analysis for Heart Failure Clinical Trial Evaluation and Interpretation. JACC: Heart Failure. Volume 8, Issue 12, December 2020, Pages 973-983. doi: 10.1016/j.jchf.2020.07.005

- Függ az időbeli cut-off megválasztásától
- Érdeemes lehet a legnagyobb szóba jövőt használni, illetve több különbözővel is ellenőrizni az eredményt

Cure modellek: a gyógyulás modellezése

- A klasszikus megközelítés nem „törődik” azzal, ha valaki meggyógyul (precízen: előbb-utóbb mindenki elszenvedi a végpontot)...
- ...egyszerűen cenzorálódni fog (miközben a cenzorálás igazából két dolgot jelentene: akik lost to follow-up, és akik *még* nem érték el a végpontot)
- Ők azonban soha nem fogják elérni
- All-cause mortalitásnál ez logikus, hiszen ott nincs is más lehetőség
- De mi van, ha okspecifikus mortalitást használunk?
- A gyógyulóknak lényegében végtelen a túlélésük (a túlélési görbe nem nullába tart, a cenzoráltak nem tudjuk, hogy hová tartoznak, gyógyulóak, vagy nem, csak szokásos – két – értelemben cenzorálódtak), ezt valahogy kezelni kellene
- A cenzoráltak hovatartozása miatti bizonytalanság következtében nem triviális a matematikai kezelés
- (Az egy külön kérdés, hogy mi van, ha más okból fakadó halálozás is van)

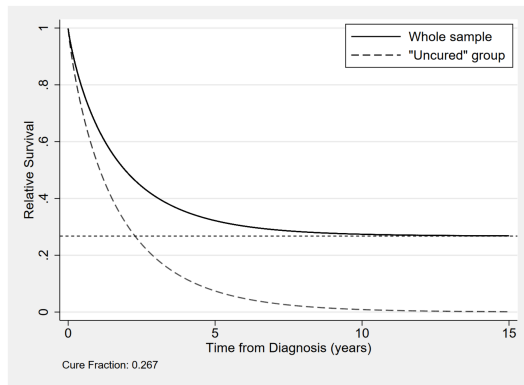
Amico, M., Van Keilegom, I. (2018). Cure Models in Survival Analysis. *Annual Review of Statistics and Its Application*, 5(1), 311–342.

doi:10.1146/annurev-statistics-031017-100101

Cure modellek: keverék (mixture) modellek

- A betegek egy adott hányada meggyógyul (100% valószínűséggel túlél), a többiek túlélését egy szokásos túlélési görbe írja le
- A túlélés adott időpontban tehát a gyógyulók aránya, plusz a nem gyógyulók közül annyi, aki még túlél
- Tehát itt két becsülendő paraméter van: a gyógyulók aránya, és a túlélési görbe a nem gyógyulók körében

Cure modellek: keverék (mixture) modellek



Paul C. Lambert. Modeling of the cure fraction in survival studies. The Stata Journal (2007) 7, Number 3, pp. 351–375. DOI: 10.1177/1536867X0700700304.

Cure modellek: nem keverék modellen alapuló (non-mixture) modellek

- Az alapötlet, hogy a hazard belefut egy felső határba...
- ...ebből fakadóan a túlélési függvény egy alsóba
- És pont ezt akartuk modellezni
- (A matematikája a proporcionális hazard modellre emlékeztet)
- Ritkábban használják

Egy általános gondolat a relatív és abszolút metrikákhoz

- Az abszolút metrika (ARR, medián túlélési különbség, 1 éves túlélési arány különbség stb.) felfogható származtatott mutatóként: a relatív metrika (RR, OR, HR stb.) és a baseline kockázat szorzata
- Bináris végpontnál egyszerű a helyzet:
$$ARR = p_C - p_T = p_C - p_C \cdot RR = p_C \cdot (1 - RR) = p_C \cdot RRR$$
 (tehát *szó szerint* szorzat)
- Bár esetről esetre meg lehetne vizsgálni, de *nagy általánosságban* az mondható, hogy
 - az abszolút metrikák fontosak a klinikai döntéshozatalhoz,
 - kutatásban viszont a relatívat érdemes lemérni, mert stabilabb, jobban átvihető más populációra (más baseline kockázatra)
- Avagy mi a baj a korábbi sztatinos példával

Smeeth L, Haines A, Ebrahim S. Numbers needed to treat derived from meta-analyses—sometimes informative, usually misleading. *BMJ*. 1999 Jun 5;318(7197):1548-51. Ferenci T.

Absolute risk reduction may depend on the duration of the follow-up. *Expert Rev Clin Pharmacol*. 2017 Dec;10(12):1409-1410.

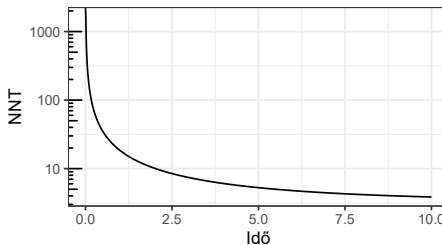
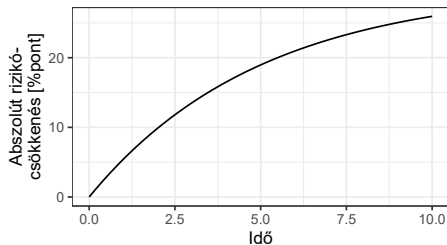
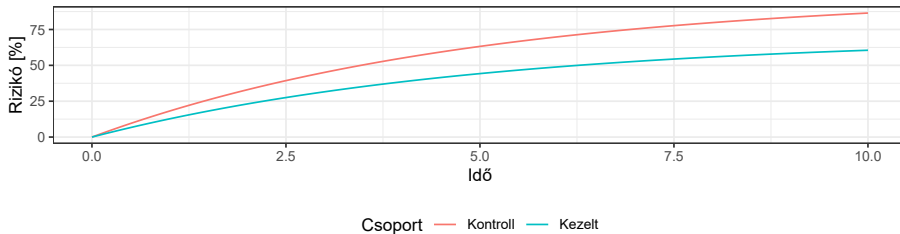
A mutató stabilitásának kérdése: a sztatínok esete

Kísérlet neve	Kontrollcsoport rizikója	Relatív rizikó	Abszolút rizikó-különbség	Utánkövetés hossza [év]
JUPITER	0,48%	0,81 (-19%)	0,09 %pont	1,9
AFCAPS/TexCAPS	0,76%	0,68 (-32%)	0,24 %pont	5,2
ASCOT-LLA	1,60%	0,90 (-10%)	0,16 %pont	3,3
WOSCOPS	2,22%	0,68 (-32%)	0,70 %pont	4,9
CARE	6,26%	0,86 (-14%)	0,87 %pont	5
HPS	9,13%	0,83 (-17%)	1,52 %pont	5
4S	9,31%	0,66 (-34%)	3,19 %pont	5,4
LIPID	9,62%	0,76 (-24%)	2,28 %pont	6,1
PROSPER	10,06%	0,86 (-14%)	1,38 %pont	3,2

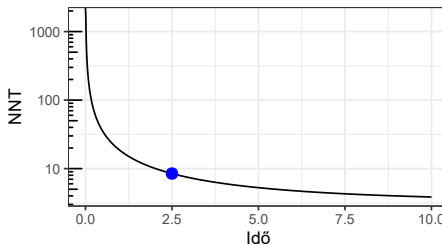
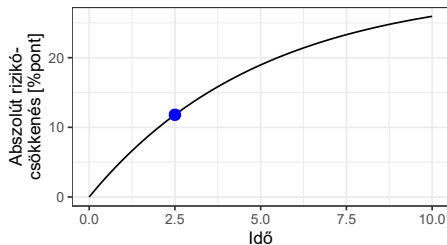
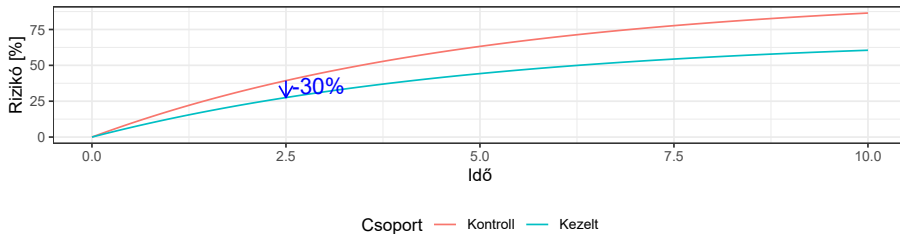
Mi volt a baj az első sztatinos példával?

- A kérdés kulcsa az utánkövetési idő: vélhetően nagyon rövid volt!
- (Mindössze 1 meg 2% kapott halálos infarktust?!)
- Itt jön a kritikus kérdés: *ha* elfogadjuk a relatív mutató stabilitását *időben is*, akkor azt látjuk, hogy – a baseline rizikó akkumulálódásával – az ARR egy nagyobb és nagyobb (és így az NNT egyre kisebb és kisebb) lesz

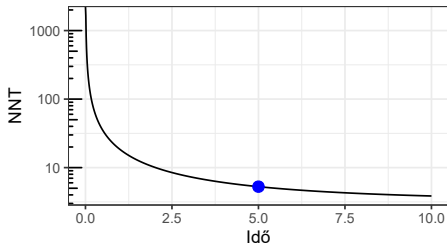
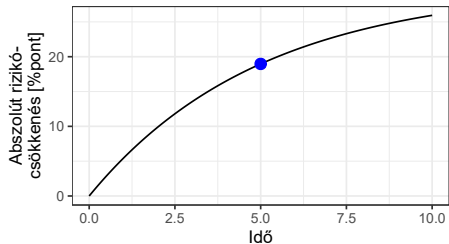
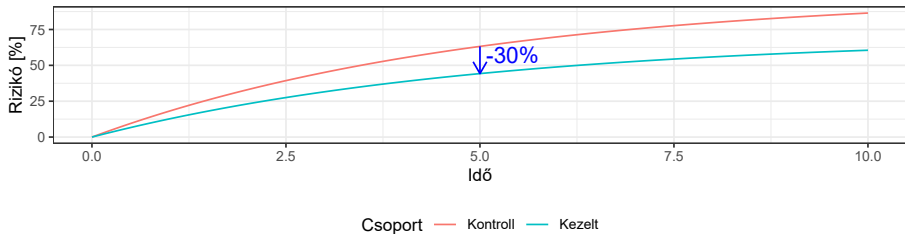
Mi történik időben, stabil RR mellett?



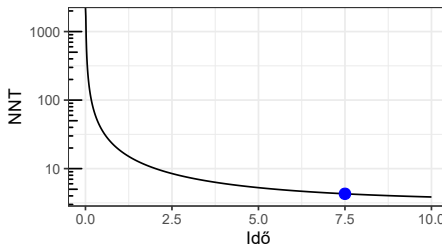
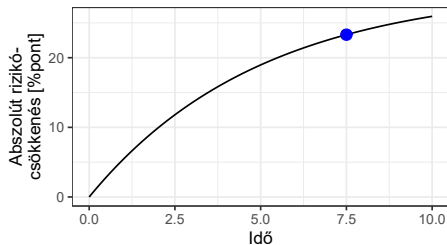
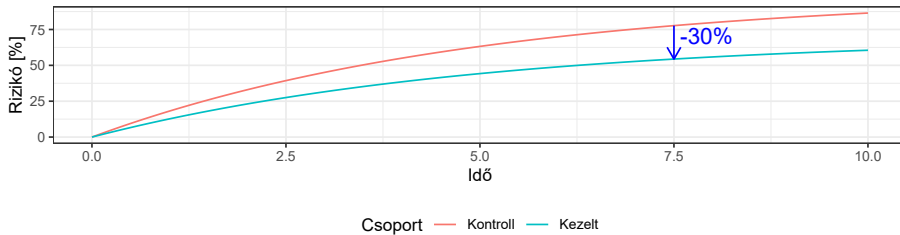
Mi történik időben, stabil RR mellett?



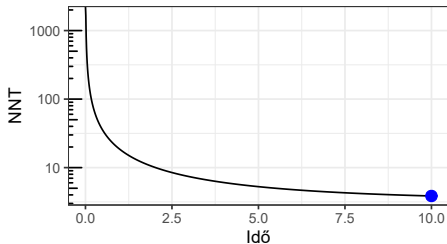
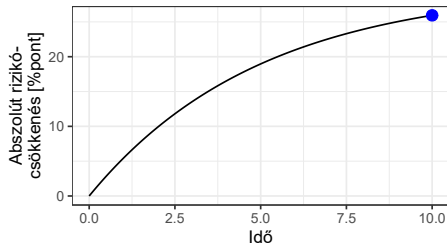
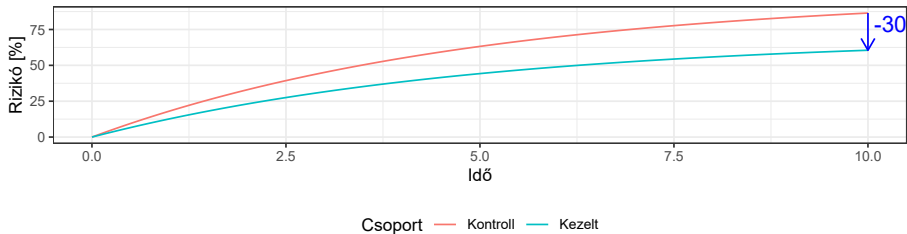
Mi történik időben, stabil RR mellett?



Mi történik időben, stabil RR mellett?



Mi történik időben, stabil RR mellett?



2. számú aspektus: Problémát jelent az RCT-k speciális betegösszetétele?

- A fenti jelenség természetesen nem csak a különböző utánkövetési idők által indukált eltérő kockázatokra igaz, hanem az eltérő betegösszetételből fakadó eltérő kockázatra is!
- Gyakran hallani: „az RCT-kben mindig speciális betegeket vonnak be, így kérdés, hogy az eredményeik mennyire általánosíthatóak az összes betegre (ami ebből fakadóan a megfigyeléses vizsgálatok előnye)”
- Ez igaz...

2. számú aspektus: Problémát jelent az RCT-k speciális betegösszetétele?

Ez igaz...

Table 3 Percentage of individuals who potentially meet eligibility criteria for included asthma randomised controlled trials

RCT reference	Current asthma (%), n = 179	Current asthma on treatment (%), n = 127
13	5	7
14	7	10
15	6	9
16	6	8
17	0	0
18	4	6
19	2	3
20	1	1
21	7	9
22	8	11
23	7	9
24	36	43
25	2	2
26	1	0
27	1	2
28	2	3
29	3	5

RCT, randomised controlled trial.

Travers J, Marsh S, Williams M, Weatherall M, Caldwell B, Shirtcliffe P, Aldington S, Beasley R. External validity of randomised controlled trials in asthma: to whom do the results of the trials apply? *Thorax*. 2007 Mar;62(3):219-23.

2. számú aspektus: Problémát jelent az RCT-k speciális betegösszetétele?

- A fenti jelenség természetesen nem csak a különböző utánkövetési idők által indukált eltérő kockázatokra igaz, hanem az eltérő betegösszetételből fakadó eltérő kockázatra is!
- Gyakran hallani: „az RCT-kben mindig speciális betegeket vonnak be, így kérdés, hogy az eredményeik mennyire általánosíthatóak az összes betegre (ami ebből fakadóan a megfigyeléses vizsgálatok előnye)”
- Ez igaz, de...
- Fogjuk fel úgy, hogy a táblázat alsó soraiban lévő vizsgálatok a „valódi” betegpopuláción történtek, a felső sorokban lévők a klinikai kísérlet alanyai!

3. számú aspektus: egyéni döntéshozatal

- Egész idáig csoportokról, átlagokról, aggregált adatokról beszéltünk
- De az egész logika változatlanul alkalmazható az egyéni döntéshozatal kontextusában is!
- „Írjak-e fel sztatint ennek a konkrét, előttem ülő páciensnek?”
- A metodika a fentiek alapján:
 - Kiveszem az RR -t a gyógyszervizsgálatból, a p_C -t pedig a beteg jellemzői alapján
 - A kettőt összevetve kiszámolom az ARR -t
 - Abból kiszámolom az NNT -t
 - És ezzel meghatározom a kockázat-haszon mérleget a gyógyszeradási döntéshez – egyedileg, az adott, konkrét betegre!
- Megjegyzés: honnan tudjuk *egyénil*eg a baseline rizikót, a p_C -t az adott, konkrét betegre? – a jellemzői alapján (életkor, nem, társbetegségek stb.); a különböző risk score-ok pont ezt csinálják

Az abszolút és relatív metrikák használata

- Tehát, az általánosan jó stratégia: a vizsgálatban a relatív metrikát mérni, de aztán a konkrét klinikai szituációban az abszolút metrika kiszámításával kontextusba helyezni azt
- A cél: „additive at the point of analysis but relevant at the point of application” (Stephen Senn)
- Általában is igaz, hogy „le kell fordítani” az eredményeket (klinikailag: az általános információból a konkrét betegre vonatkozóra, figyelembe véve az egyéni információkat; matematikailag: RR-ről ARR-re, figyelembe véve a p_C -t)
- A fenti munkamódszer tehát lehetővé teszi az extrapolációt, időben, vagy különböző páciensek között, de ez igazából ugyanazt jelenti: különböző baseline rizikókra áttérés
- De, mindez úgy működött, hogy *feltételeztük*, hogy a relatív rizikó állandó
- Ez egy feltételezés (ezért nem lenne rossz például, ha az RCT-knek nem lenne annyira speciális a betegösszetétele, hogy ne nagyon legyen szükség ilyen feltevésekre)

Senn S. At odds with reality. <http://www.bmj.com/rapid-response/2011/10/27/odds-reality>

Egy illusztráció az abszolút és relatív metrikákhoz...

ICS-LAMA-LABA vs. LAMA-LABA COPD-ben

The rate of moderate or severe exacerbations during treatment among patients assigned to triple therapy was 0.91 per year, as compared with [...] 1.21 per year among those assigned to umeclidinium–vilanterol (rate ratio with triple therapy, 0.75; 95% CI, 0.70 to 0.81; 25% difference; $P < 0.001$).

- Szokásos sztori: a relatív („25%-os csökkenés”) jó lehet mint kísérletben kimérendő eredménye...
- ...de a klinikai döntéshozatalhoz az abszolút lesz a fontos!
- A másik fontos: a kontextusba helyezés, például a mi betegünknek a baseline kockázata is nagyobb (a vizsgálatban egészségesebb alanyok voltak!), mondjuk 1,6 exacerbáció/év
- *Feltéve*, hogy a gyógyszer relatív hatása állandó, a tripla kombinációval $1,6 \cdot 0,75 = 1,2$ exacerbációja lesz évente

Lipson DA, Barnhart F, Brealey N, et al. Once-Daily Single-Inhaler Triple versus Dual Therapy in Patients with COPD. N Engl J Med. 2018 May 3;378(18):1671-1680.

Egy illusztráció az abszolút és relatív metrikákhoz...

- A különbség $(1,6 - 1,2 =) 0,4$ exacerbáció/év
- Még kifejezőbb ha úgy fogalmazunk: $(1/0,4 =) 2,5$ évig kell egy beteget kezelnünk ezzel ahhoz, hogy egy exacerbációt megelőzzünk
- Biztos megéri 2,5 évnyi kezelés mellékhatásainak kitenni egy beteget egy exacerbáció megelőzése végett? Biztos megéri 2,5 évnyi kezelést kifizettetni egy exacerbáció megelőzése végett?
- Ezek persze már nem statisztikai kérdések, de a „statisztika” segít abban, hogy jól megfogalmazzuk/megértsük, hogy egyáltalán mi a kérdés!

Lipson DA, Barnhart F, Brealey N, et al. Once-Daily Single-Inhaler Triple versus Dual Therapy in Patients with COPD. N Engl J Med. 2018 May 3;378(18):1671-1680.

...és egy – nagyon fontos – továbbgondolás

- A másik nagy előny ez abszolút metrikáknak, hogy megkönnyíti a különböző kimenetek egymással történő összevetését

ICS-LAMA-LABA vs. LAMA-LABA COPD-ben

[T]he risk of clinician-diagnosed pneumonia was significantly higher with triple therapy than with umeclidinium-vilanterol, as assessed in a time-to-first event analysis (hazard ratio, 1.53; 95% CI, 1.22 to 1.92; $P < 0.001$).

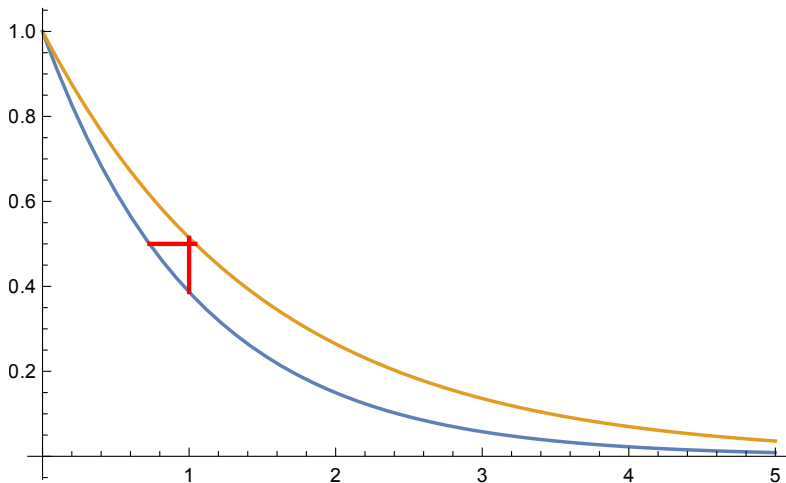
- A 25% exacerbáció rizikó csökkenés *összevethetetlen* az 53%-os pneumoniabeli növekedéssel! (hiszen teljesen mások a baseline rizikók)
- Ha azonban jobban megnézzük a kéziratot, akkor láthatjuk, hogy az ICS-LAMA-LABA csoportban a pneumonia rizikója 0,0958 / betegév, a LAMA-LABA csoportban 0,0612 / betegév
- Ez lehetővé teszi, hogy abszolút metrikát számoljunk: a tripla kombináció $(0,0958 - 0,0612 =) 0,0346$ /évvel növeli a pneumonia rizikóját (mondjuk, hogy itt a betegünk baseline rizikója ugyanaz, mint a kutatásban)
- Azaz a szokásos további átszámítással: $(1/0,0346 =) 28,9$ évnyi kezeléssel okozunk egy többlet tüdőgyulladást

- És most már ordít, hogy miért beszélhetünk „összevethetőség megteremtéséről”: ez magyarul azt jelenti, hogy kb. 11,5 exacerbációt előzünk meg 1 tüdőgyulladás árán!
- Megint csak: az nyilván nem statisztikai kérdés, hogy ez megéri-e, „ér-e” 11,5 exacerbációt 1 pneumonia – de így legalább látjuk, mégpedig klinikailag értelmes, releváns módon, hogy egyáltalán mit kell összevetni! (a százalékos változásokból ez nem derült ki!)
- (Megjegyzés: út az exacerbáció és a pneumonia közös nevezőre hozására: életév-veszteség, minőséggel korrigált életév-veszteség...)

Sajnos az abszolút mutatók sem problémamentesek

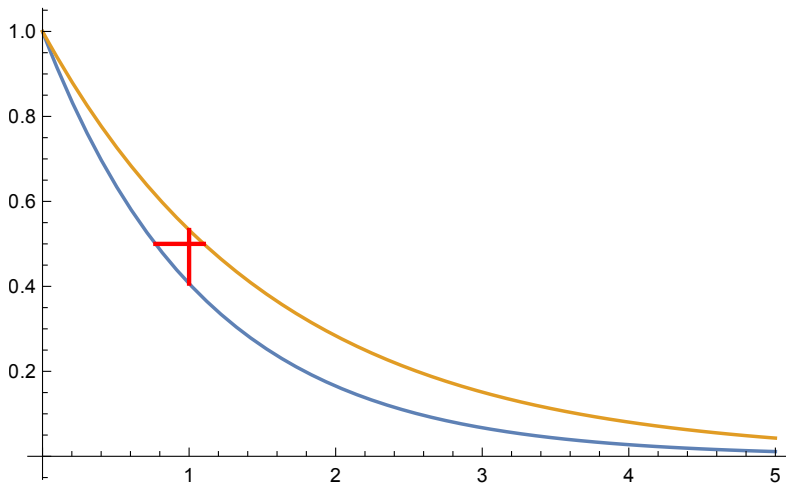
- Eltelt idő jellegű végpont esetén a két abszolút jellegű mutató (medián túlélés növekedés és fix – például 1 éves – túlélési arány javulás) sajnos *ellentmond egymásnak*!
- A következő ábrákon pontosan ugyanolyan a túlélési görbe alakja, pontosan ugyanúgy rögzítetten 0,7 a HR, az egyetlen különbség, hogy a kontrollcsoport milyen gyorsan hal (mekkora a baseline rizikó)

Sajnos az abszolút mutatók sem problémamentesek



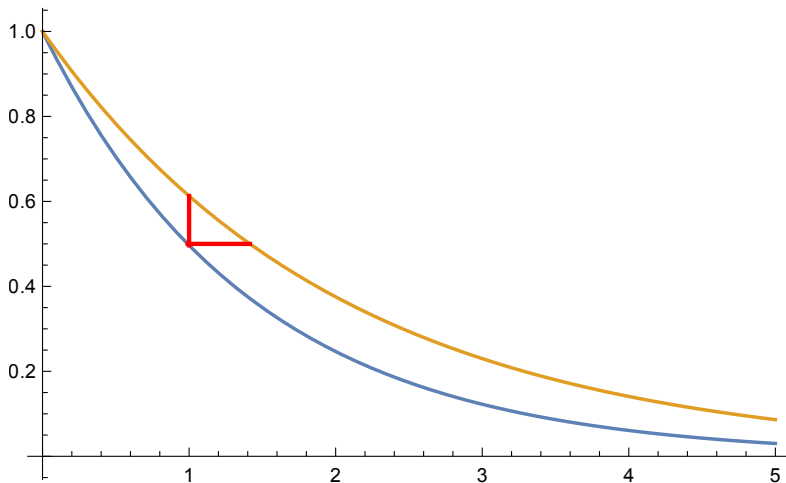
Snappinn S, Jiang Q. On the clinical meaningfulness of a treatment's effect on a time-to-event variable. Stat Med. 2011 Aug 30;30(19):2341-8.

Sajnos az abszolút mutatók sem problémamentesek



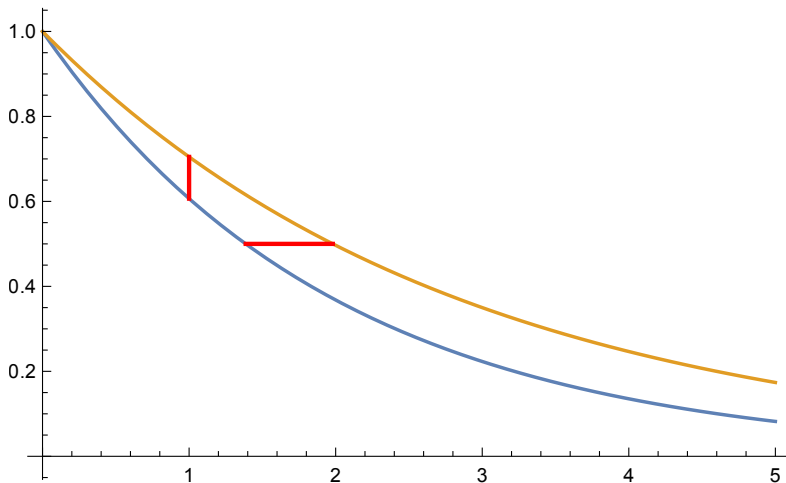
Snappinn S, Jiang Q. On the clinical meaningfulness of a treatment's effect on a time-to-event variable. Stat Med. 2011 Aug 30;30(19):2341-8.

Sajnos az abszolút mutatók sem problémamentesek



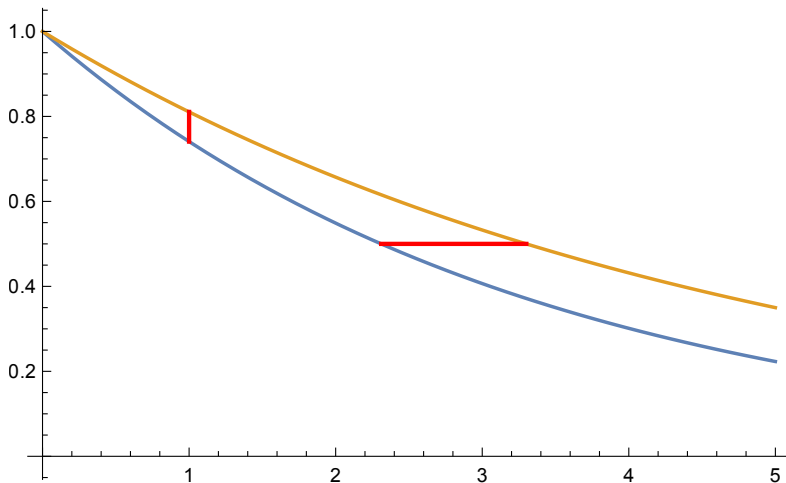
Snappinn S, Jiang Q. On the clinical meaningfulness of a treatment's effect on a time-to-event variable. Stat Med. 2011 Aug 30;30(19):2341-8.

Sajnos az abszolút mutatók sem problémamentesek



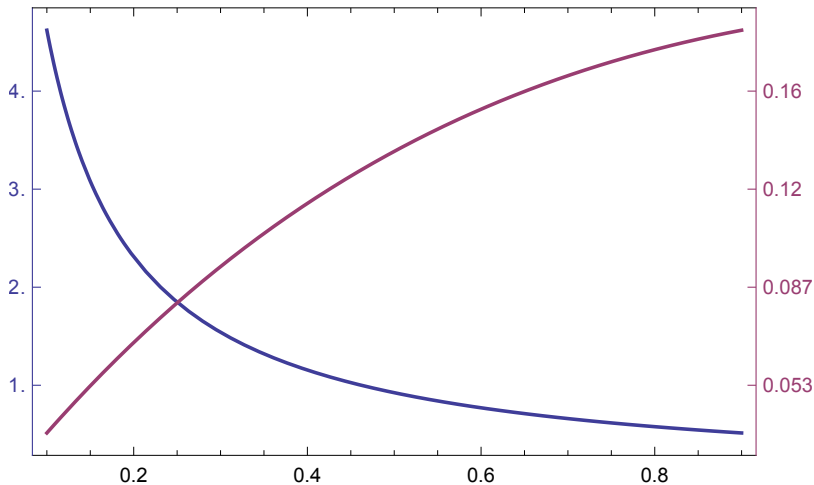
Snappin S, Jiang Q. On the clinical meaningfulness of a treatment's effect on a time-to-event variable. Stat Med. 2011 Aug 30;30(19):2341-8.

Sajnos az abszolút mutatók sem problémamentesek



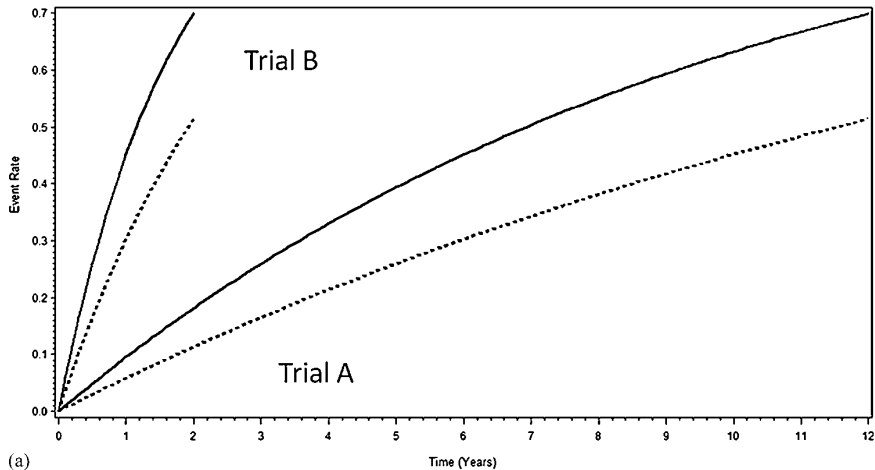
Snappinn S, Jiang Q. On the clinical meaningfulness of a treatment's effect on a time-to-event variable. Stat Med. 2011 Aug 30;30(19):2341-8.

A két mutató ellentmondása



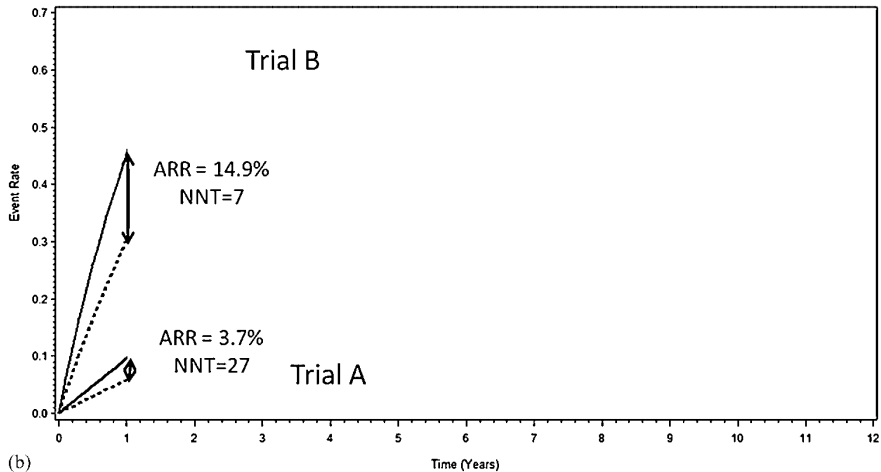
Snapinn S, Jiang Q. On the clinical meaningfulness of a treatment's effect on a time-to-event variable. Stat Med. 2011 Aug 30;30(19):2341-8.

Az ellentmondás egy konkrét illusztrációja



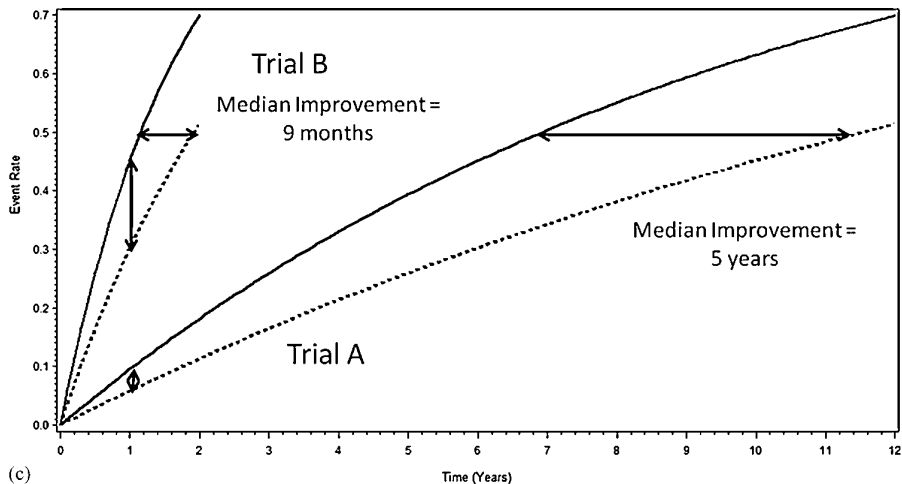
Snapinn S, Jiang Q. On the clinical meaningfulness of a treatment's effect on a time-to-event variable. Stat Med. 2011 Aug 30;30(19):2341-8.

Az ellentmondás egy konkrét illusztrációja



Snapinn S, Jiang Q. On the clinical meaningfulness of a treatment's effect on a time-to-event variable. Stat Med. 2011 Aug 30;30(19):2341-8.

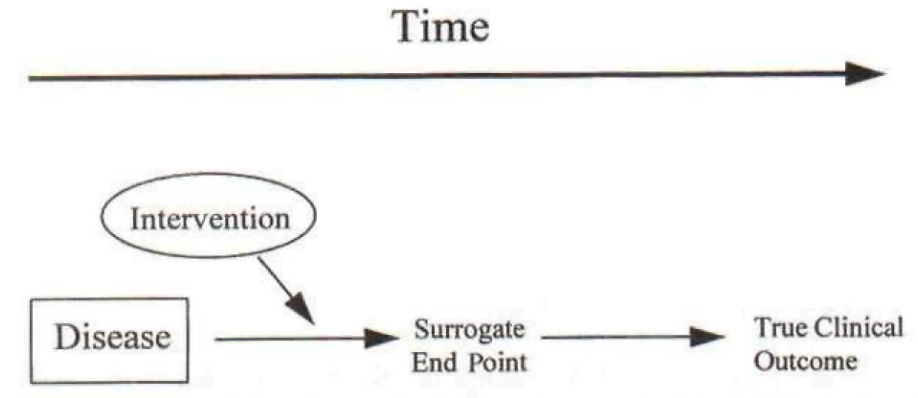
Az ellentmondás egy konkrét illusztrációja



(c)

Snapinn S, Jiang Q. On the clinical meaningfulness of a treatment's effect on a time-to-event variable. Stat Med. 2011 Aug 30;30(19):2341-8.

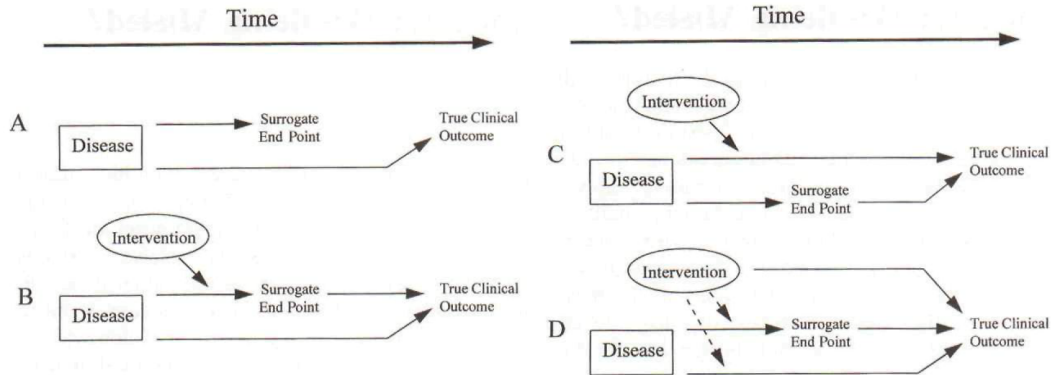
Kemény vs. surrogate végpont: ami az optimális



Tételmondat: a jó surrogate-séghez *nem* elég, hogy jól korrelált legyen („correlate does not a surrogate make”)!

Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? Ann Intern Med. 1996 Oct 1;125(7):605-13.

Kemény vs. surrogate végpont: ami baj lehet



Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? Ann Intern Med. 1996 Oct 1;125(7):605-13.

Kemény vs. surrogate végpont: pár példa

Table 1. Speculation on Reasons for Failures of Surrogate End Points*

Disease and Intervention	End Points		Settings in Figure 1†			
	Surrogate	Clinical	A	B	C	D
Cardiologic disorder						
Arrhythmia						
Encainide; flecainide	Ventricular arrhythmias	Survival		+		++
Quinidine; lidocaine	Atrial fibrillation	Survival		+		++
Congestive heart failure						
Milrinone; flosequinan	Cardiac output; ejection fraction	Survival		+		++
Elevated lipid levels						
Fibrates; hormones; diet; lovastatin	Cholesterol levels	Survival		+		++
Elevated blood pressure						
Calcium channel blockers	Blood pressure	Myocardial infarction; survival		+		++
Cancer						
Prevention						
Finasteride	Prostate biopsy	Symptoms; survival	+++‡			
Advanced disease						
Fluorouracil plus leucovorin	Tumor shrinkage	Survival		+		++
Other diseases						
HIV infection or AIDS						
Antiretroviral agents	CD4 levels; viral load	AIDS events; survival		+	+	+
Osteoporosis						
Sodium fluoride	Bone mineral density	Bone fractures	+			+
Chronic granulomatous disease						
Interferon-γ	Bacterial killing; superoxide production	Serious infection			++	

* AIDS = acquired immunodeficiency syndrome; HIV = human immunodeficiency virus; + = likely or plausible; ++ = very likely.

† A = surrogate end point not in causal pathway of the disease process; B = of several causal pathways of the disease, the intervention only affects the pathway mediated through the surrogate; C = the surrogate is not in the pathway of the intervention's effect or is insensitive to its effect; D = the intervention has mechanisms of action that are independent of the disease process.

‡ In settings in which only latent disease is prevented.

Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? Ann Intern Med. 1996 Oct 1;125(7):605-13.

- A legfontosabb ok: dúsítja a végpontok számát, ezáltal növeli az erőt (adott mintanagysággal kisebb különbséget is ki tudunk mutatni, illetve ugyanazon különbség kimutatásához elég kisebb mintanagyság is)
- Védelem az ellen, ha a terápia csak konvertálja a végpontot
- De: ugyanazon biológiai jelenség manifesztációit mérjük, de azért ne túlságosan ugyanazt
- Mely komponens változása adta az összesített változást...?
- Extrém széles összefogás kérdései (pl. UKPDS)

Egy példa: UKPDS

	Active rx. (n = 2729)	Conv. rx. (n = 1138)	P-value	Rel. risk	95% CI*
Composite endpoints	963	238	0.029	0.88	[0.79 – 0.99]
Fatal endpoints	285	129	0.340	0.90	[0.73 – 1.11]
Fatal MI	207	90	0.630	0.94	[0.68 – 1.30]
Stroke deaths	43	15	0.600	1.17	[0.54 – 2.54]
Renal deaths	8	2	0.530	1.83	[0.21 – 12.49]
Glucose related**	1	1	0.523	0.420***	[0.03 – 6.66]
Sudden death	24	18	0.047	0.54	[0.24 – 1.21]
Death from PVD****	2	3	0.120	0.26	[0.03 – 2.77]
Nonfatal endpoints					
Nonfatal MI	197	101	0.067	0.79	[0.58 – 1.09]
Angina pectoris	177	72	0.910	1.02	[0.71 – 1.46]
Major stroke	114	44	0.720	1.07	[0.68 – 1.69]
Amputation	27	18	0.990	0.81	[0.28 – 1.33]
Blindness	78	38	0.390	0.84	[0.51 – 1.40]
Renal failure	16	9	0.450	0.76	[0.53 – 1.08]
Photocoagulation*****	207	117	0.003	0.71	[0.53 – 0.96]

UKPDS Study Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group. Lancet. 1998 Sep 12;352(9131):837-53.