

A confounding problémája

Ferenci Tamás
tamas.ferenci@medstat.hu

2017. október 22.

A confounding problémája

- Naiv gondolat arra, hogy hogyan vizsgáljuk a kérdéseket: több rákos van a távvezeték közelében élők körében? a több vörös húst evők körében? több T1DM-es van a császármetszéssel születők körében?
- Roppant csábító, de... *teljesen fals!*
- A távvezeték körében élő emberek nem *csak* a távvezeték közelében élésben térnek el a nem távvezeték közelében élőktől, a több vörös húst evők nem *csak* a több vörös hús evésében térnek el a kevesebb vörös húst evőktől, a császármetszéssel születők nem *csak* a császármetszéssel születésben térnek el a pvn születőktől stb. stb.

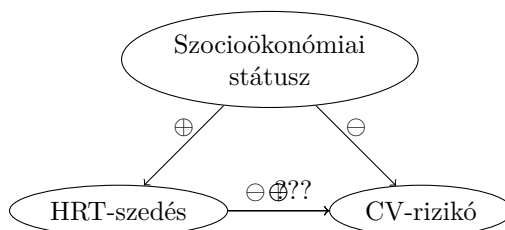
Greenland S, Morgenstern H. Confounding in health research. *Annu Rev Public Health*. 2001;22:189-212.

A confounding problémája

- Innen kezdve, ha találunk is különbséget a végpontban, nem tudhatjuk, hogy az mi miatt van: a *vizsgált* eltérés (azaz az expozíció) miatt, a vizsgált eltéréssel *együtt járó* valamely *egyéb eltérés* miatt, vagy a kettő valamilyen keveréke miatt
- Ha a két csoport nem *csak* az expozícióban tér el, onnantól biztosan *nem* tudhatjuk, hogy ha találunk is különbséget, annak mi az oka!
- (Az ideális kutatási módszerre visszautalva: az a gond, hogy az összehasonlító csoportra nem igaz, hogy mutatja mi történne az exponált csoportban, ha nem lett volna expozíció...)

Greenland S, Morgenstern H. Confounding in health research. *Annu Rev Public Health*. 2001;22:189-212.

A confounding problémája



Nézzünk egy számszerű példát is erre

CV események előfordulása:

| | Nem szed HRT-t | Szed HRT-t |
|----------------------------------|------------------|----------------|
| Alacsony szocioökonómiai státusz | 4% (240/6000) | 6% (12/200) |
| Magas szocioökonómiai státusz | 1% (50/5000) | 2% (60/3000) |
| Összességében | 2,6% (290/11000) | 2,3% (72/3200) |

A confounder definíciója tehát

Változó, amire egyszerre igaz, hogy:

1. összefügg az expozícióval (tehát eltérnek e szerint az exponált és nem exponált csoportok), és
2. önmagában is hat a végpontra

A HRT-s példában:

1. A szocioökonómiai státusz összefügg a gyógyszereszedéssel (mert alacsonyban csak $200/6200 = 3,23\%$ szed HRT-t, a magasban viszont $5000/8000 = 62,5\%$)
2. A szocioökonómiai státusz önmagában is hat a végpontra (1%-ról 4%-ra, illetve 2%-ról 6%-ra emeli a kockázatot, ha alacsony)

Nincs confounding, ha a változó nem függ össze az expozícióval

| | Nem szed HRT-t | Szed HRT-t |
|----------------------------------|------------------|-----------------|
| Alacsony szocioökonómiai státusz | 4% (186/4650) | 6% (93/1550) |
| Magas szocioökonómiai státusz | 1% (60/6000) | 2% (40/2000) |
| Összességében | 2,3% (246/10650) | 3,7% (133/3550) |

(Mindkét csoportban $1550/6200 = 6000/8000 = 25\%$ a szedők aránya)

Nincs confounding, ha a változó nem hat a végpontra

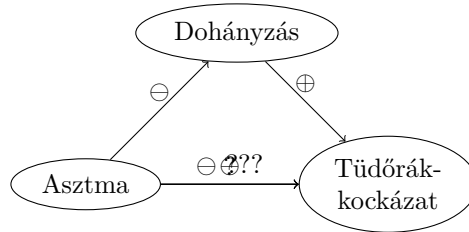
| | Nem szed HRT-t | Szed HRT-t |
|----------------------------------|----------------|---------------|
| Alacsony szocioökonómiai státusz | 3% (180/6000) | 5% (10/200) |
| Magas szocioökonómiai státusz | 3% (150/5000) | 5% (150/3000) |
| Összességében | 3% (330/11000) | 5% (160/3200) |

Két megjegyzés

- Megoldja a confounding problémáját, ha rájövünk, hogy mi a confounder, és a szerint „lebontva” (rétegezve) nézzük az eredményeket – erre a gondolatra később még visszatérünk
- Nem kötelező, hogy csak két (exponált és nem exponált) csoport legyen, sőt, általában a kauzalitási következtetés robusztusabb is, ha nem egyszerűen azt tudjuk kimutatni, hogy az exponáltaknál nagyobb a kockázat, hanem azt is, hogy az egyre jobban (hosszabban, tovább stb.) exponáltaknál egyre nagyobb a kockázat

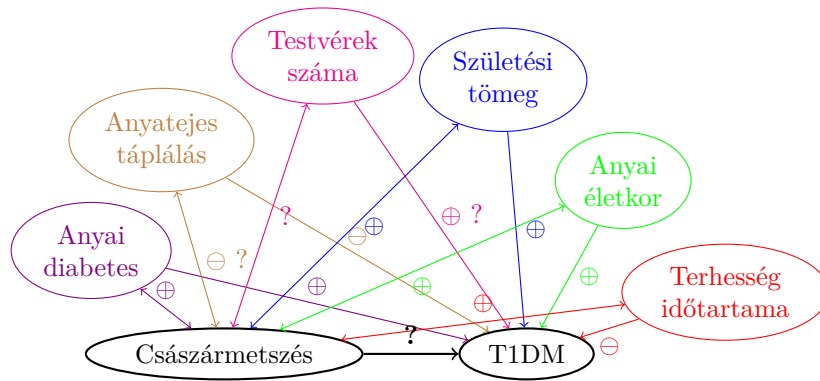
A confounding problémája

Amikor azt mondtuk a confounding első feltételeként, hogy a változó összefügg az expozícióval, akkor mindegy, hogy mi hat mire és hogyan (egyik a másikra, másik az egyikre, harmadik mindkettőre stb.), csak az a tény számít, hogy korreláltak



Pontosan emiatt a továbbiakban erre az első feltételre duplanyilas jelzést fogunk használni

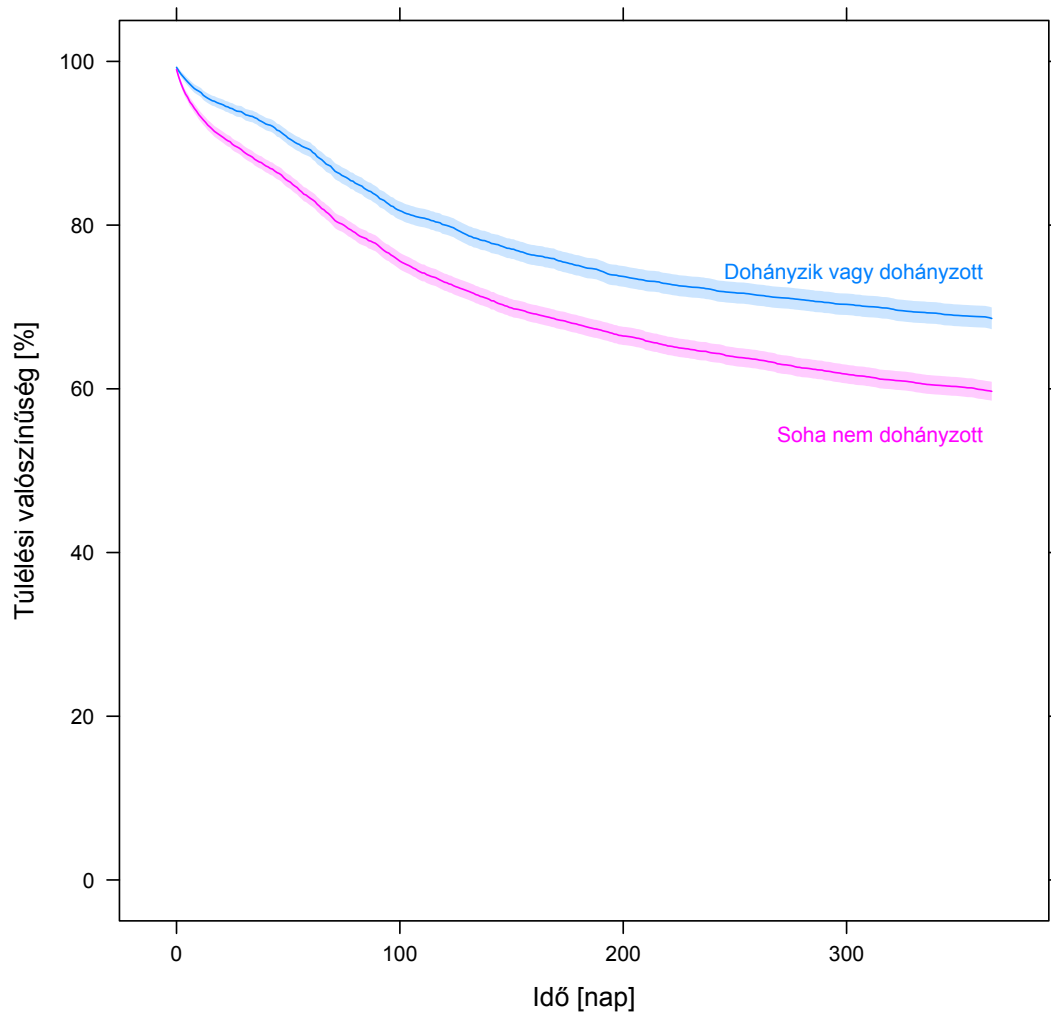
A confounding problémája



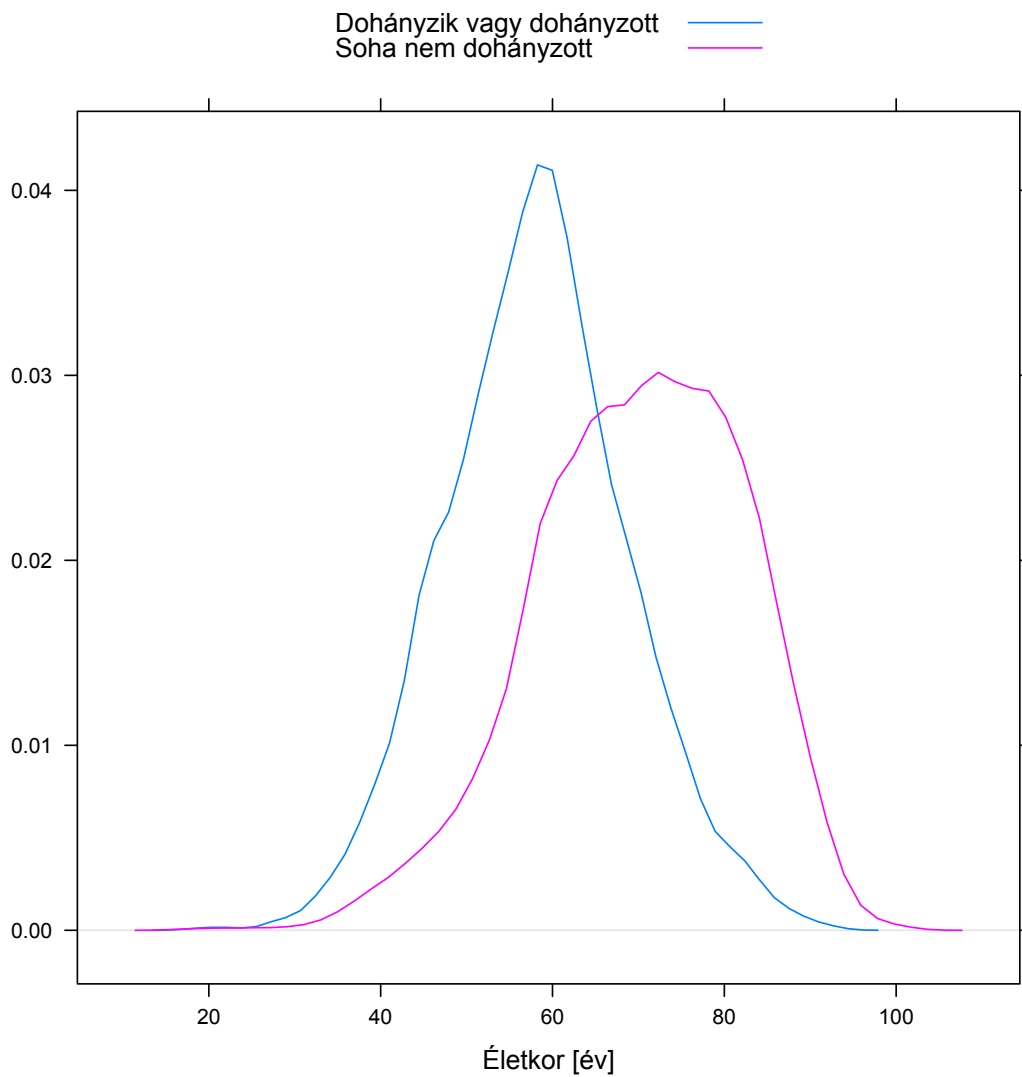
Cardwell CR, Stene LC, Joner G, et al. Caesarean section is associated with an increased risk of childhood-onset type 1 diabetes mellitus: a meta-analysis of observational studies. Diabetologia. 2008 May;51(5):726-35.

Egy klinikaibb példa a confoundingra: infarktus utáni halálozás és a dohányzás

Túlélés dohányzási státusz szerint



Egy klinikaibb példa a confoundingra: infarktus utáni halálozás és a dohányzás



Aune E, R?islien J, Mathisen M, et al. The "smoker's paradox" in patients with acute coronary syndrome: a systematic review. BMC Med. 2011 Aug 23;9:97.

A Simpson-paradoxon mint confounding

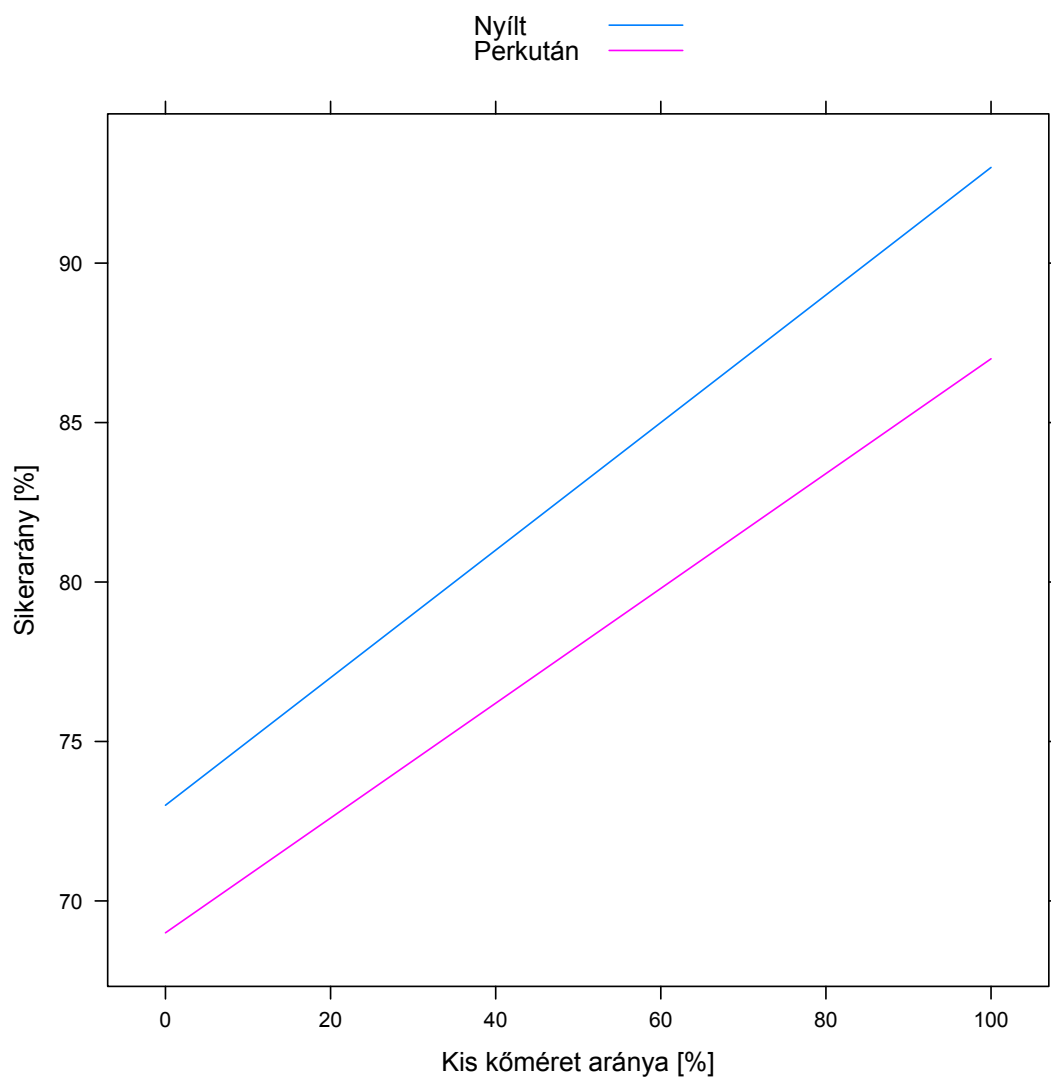
| | Nyílt feltárás | Perkután eljárás |
|---------------|----------------|------------------|
| Összességében | 78% (273/350) | 83% (289/350) |

A Simpson-paradoxon mint confounding

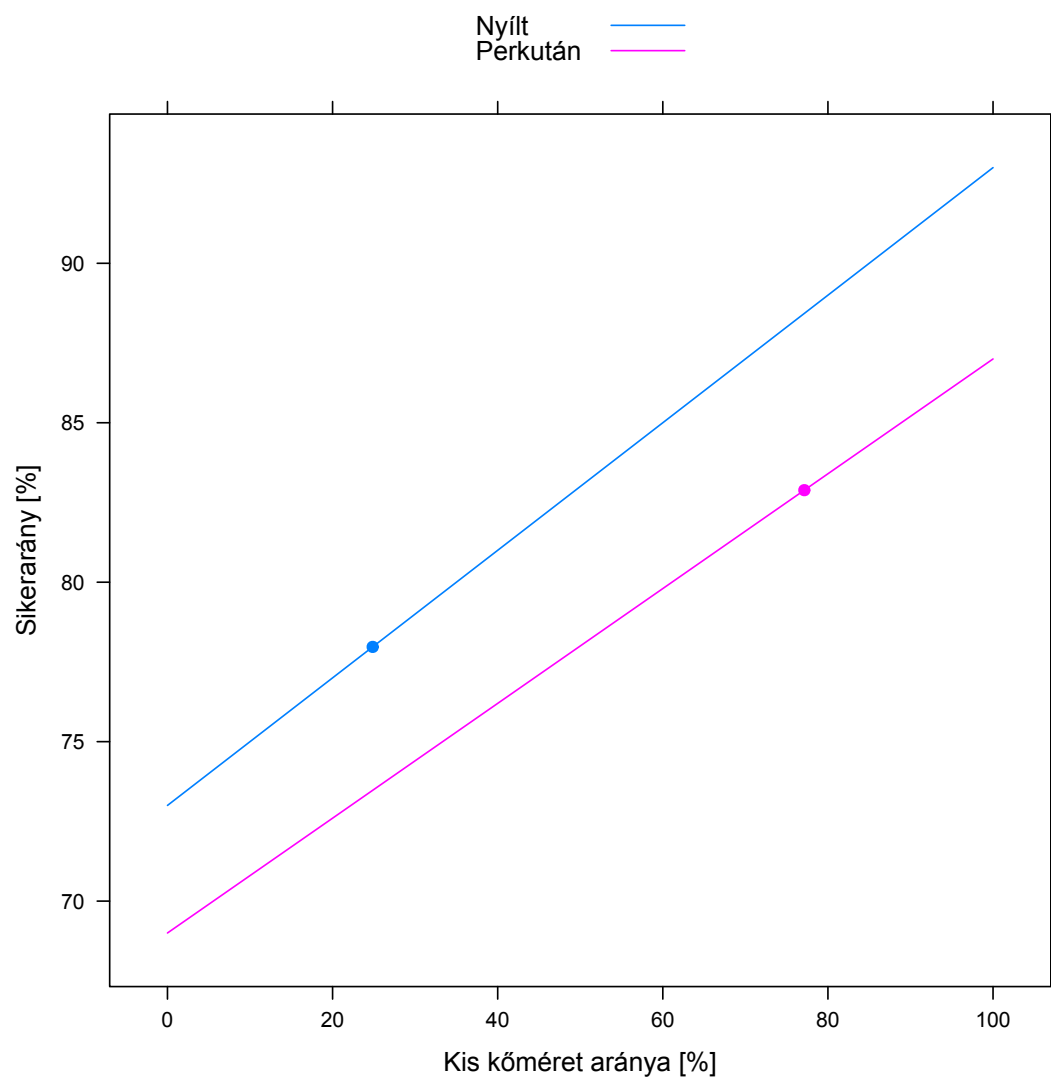
| | Nyílt feltárás | Perkután eljárás |
|----------------|----------------|------------------|
| Kőátmérő < 2cm | 93% (81/87) | 87% (234/270) |
| Kőátmérő ≥ 2cm | 73% (192/263) | 69% (55/80) |
| Összességében | 78% (273/350) | 83% (289/350) |

Julious SA, Mullee MA. Confounding and Simpson's paradox. *BMJ*. 1994 Dec 3;309(6967):1480-1. Hernán MA, Clayton D, Keiding N. *Int J Epidemiol*. 2011 Jun;40(3):780-5. Baker SG, Kramer BS. Good for women, good for men, bad for people: Simpson's paradox and the importance of sex-specific analysis in observational studies. *J Womens Health Gend Based Med*. 2001 Nov;10(9):867-72. Wainer H. The BK-Plot: Making Simpson's Paradox Clear to the Masses. *Chance*. 2002;15(3):60-62. Baker SG, Kramer BS. The transitive fallacy for randomized trials: if A bests B and B bests C in separate trials, is A better than C? *BMC Med Res Methodol*. 2002 Nov 13;2:13.

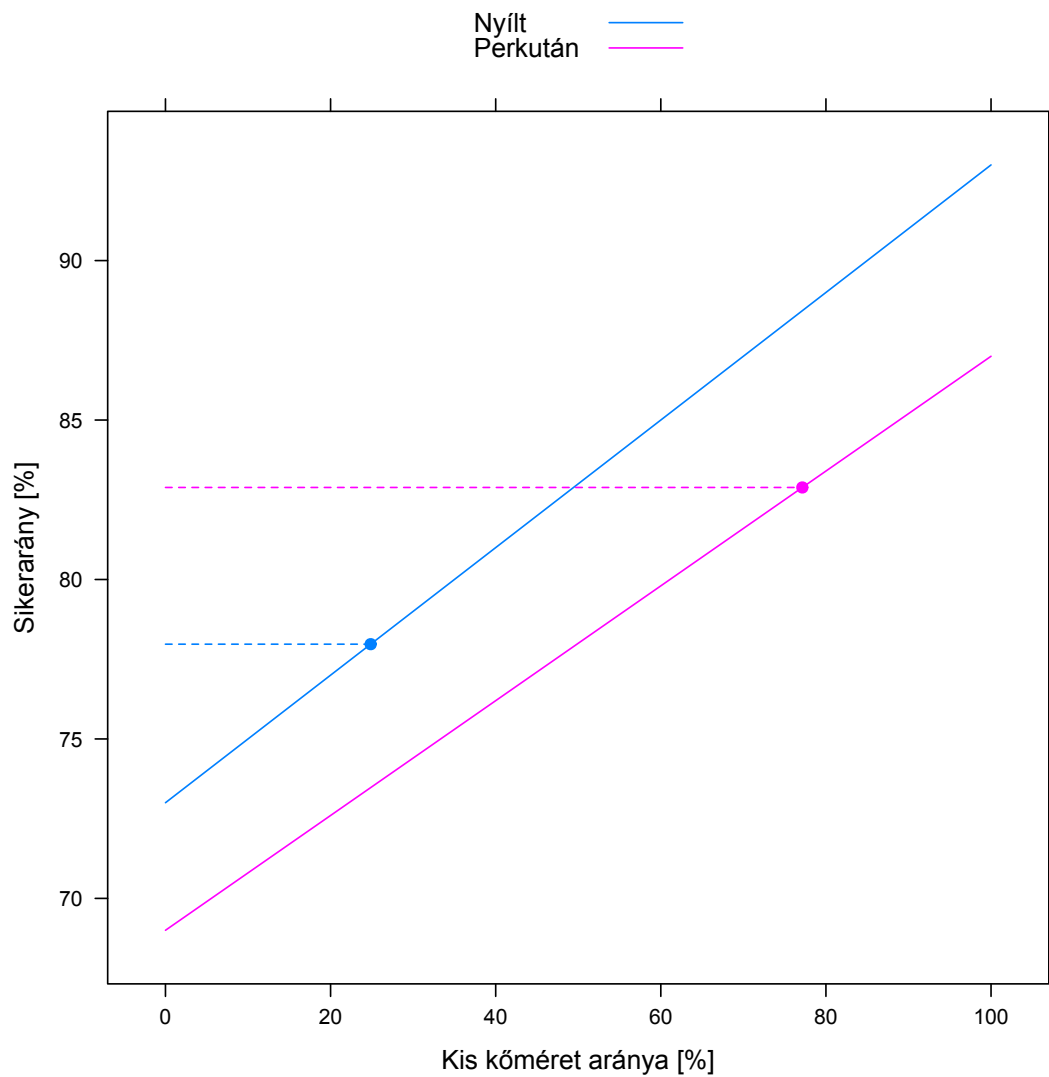
A Simpson-paradoxon grafikus magyarázata: BK-ábra



A Simpson-paradoxon grafikus magyarázata: BK-ábra



A Simpson-paradoxon grafikus magyarázata: BK-ábra



Egy nagyon más(nak tűnő) példa, ami mégis confounding

- Svédországban 2015-ben 91 ezer 709 halálozás történt, a lakosságszám 9 millió 10 ezer 729, így a CDR 10,2/ezer fő/év
- Chilében ugyanabban az évben 86 ezer 100 halálozás történt, a lakosságszám 15 millió 519 ezer 347, így a CDR 5,5/ezer fő/éve
- Svédországban kétszer (???) nagyobb a halandóság?
- ???

Egy nagyon más(nak tűnő) példa, ami mégis confounding

