A kórházi fertőzések adatainak értelmezése

Buktatók és megoldási lehetőségek

Ferenci Tamás*

2026. szeptember 2.

A magyar népegészségügy hatalmas ugrást hajt végre 2026 őszén a transzparencia terén: megkezdődik a kórházi fertőzések[1] intézmény-szintű adatainak nyilvános közlése. Az ilyen jellegű adatközlés rengeteg előnnyel bír (a jó példák és a javítandó területek azonosításával nagyban támogatja a kórházi fertőzések elleni küzdelmet, elősegíti a minőségbiztosítást, növeli az egészségügyi ellátórendszerbe vetett bizalmat), de egyúttal kihívás is, mert az adatok értelmezése sok esetben nem nyilvánvaló, miközben a hibás értelmezések komoly félreértések forrásai lehetnek.

E problémák közül is a talán legfontosabb a különböző kórházak egymással történő összehasonlítása: mennyire lehet a kórházi fertőzések előfordulása alapján összehasonlítani kórházakat? E látszólag nagyon egyszerű és nagyon kézenfekvő kérdés megválaszolása rengeteg, egyáltalán nem egyszerű és kézenfekvő problémát nyit ki.

A hazai adatközlés igyekszik reflektálni erre a kérdéskörre, és egy olyan mutatót is közöl, melynek célja kifejezetten ezen összehasonlítások elősegítése. Már most fontos azonban hangsúlyozni, hogy még ezen korrigált mutató sem tekinthető a kórházak teljesítményének valamiféle kőbevésett, biztos mérőszámának. De ez nem probléma, mert nem is ez a célja: az eredményeket egyfajta „anomáliadetekciónak" kell tekinteni, ami megjelöl bizonyos kórházakat, és pusztán annyit mond, hogy itt érdemes alaposabban vizsgálódni, hogy megértsük az eltérő teljesítmény okát. Akár még az is kiderülhet, hogy egy nagyon jónak tűnő kórháznál problémák vannak (például mert nem elég alapos a kórházi fertőzések felderítése), vagy fordítva, egy rossznak tűnő kórház példaértékű (például mert kiváló a jelentési fegyelme). Ha ez történik, az ugyanúgy azt jelenti, hogy a módszer működött és betöltötte a célját.

*Óbudai Egyetem, Élettani Szabályozások Kutatóközpont; Budapesti Corvinus Egyetem, Statisztika Tanszék. E-mail: ferenci.tamas@nik.uni-obuda.hu.

[1] A kórházi fertőzések hivatalos elnevezése egészségügyi ellátással összefüggő fertőzés. Jelen dolgozat az egyszerűbb, hétköznapi elnevezést fogja használni a könnyebb olvashatóság kedvéért, jelezve, hogy a kettő szinonima.

Jelen írás[2] célja kettős: egyrészt az alapprobléma és lehetséges megoldásainak részletes bemutatása, másrészt a hazai módszertan ismertetése. A dolgozat minden statisztikai és népegészségügyi előismeret nélkül is követhető, célja az, hogy bárki számára hozzáférhető, közérthető bevezetést adjon ezekbe a kérdésekbe. Az előbb említett problémák megoldása ugyanis nem az, hogy nem közlünk adatokat, hanem az, ha megfelelő magyarázattal *ellátva*[3] közöljük azokat. Jelen írás ezt a célt igyekszik szolgálni.

1. Bevezetés és problémafelvetés

Képzeljük el, hogy a következőt olvassuk egy újsághírben: X kórházban a kórházi fertőzések előfordulása egy adott évben 6% volt (tehát az ott kezelt betegek 6%-a kapott el kórházi fertőzést: 10 ezer beteget láttak el, és közülük 600 kapott kórházi fertőzést), míg Y kórházban ugyanez az arány 4% (3000 betegből 120). A kérdés: hol jobb a helyzet kórházi fertőzések előfordulását tekintve?

Első ránézésre a kérdés szinte érthetetlen: oda van írva a válasz, mit kérdezünk egyáltalán? Hogy a 4 kisebb-e, mint a 6…?

A dolog azonban cseles. Nézzük meg ugyanis az adatokat úgy is, ha lebontjuk aszerint, hogy mekkora volt a betegek kockázata! A „kockázat” alatt most olyan tényezőkre gondoljunk, mint a beteg életkora, társbetegségei, állapotának a súlyossága, egyáltalán, az alapbetegsége, ami sok esetben meghatározza, hogy milyen beavatkozásokat kell rajta végezni stb.[4], egyszóval, hogy a beteg mennyire fogékony arra, hogy kórházi fertőzést kapjon az őt ápoló kórháztól független, *nem* a kórház befolyása alatt álló tényezők miatt. Ezeket a jellemzőket szokták összefoglalóan betegösszetételnek is nevezni. A következő táblázat minden cellájában szerepel a megfertőződött betegek aránya, utána zárójelben, hogy ez hogyan jött ki, tehát, hogy hány beteget látott el a kórház az adott kategóriában és hogy közülük hány kapott el kórházi fertőzést:

	X kórház	Y kórház
Kis kockázatú betegek	2 % (100/ 5 000)	3 % (84/2 800)
Nagy kockázatú betegek	10 % (500/ 5 000)	18 % (36/ 200)
Összességében	6 % (600/10 000)	4 % (120/3 000)

[2] A bemutatott módszer kidolgozása és hazai viszonyokra történő adaptálása Dr. Surján Györggyel (NNGYK Egészségmonitorozási Főosztály) közösen történt. A kéziratok gondos elolvasásáért, a felmerült ötletek önzetlen és lelkes véleményezéséért Dr. Kovács Lászlónak (BCE Statisztika Tanszék) tartozok köszönettel. E dolgozat azonban teljes egészében az én munkám, így minden benne maradt hibáért engem terhel a felelősség.

[3] Ne legyenek illúzióink, még ekkor is elkerülhetetlenül elő fognak fordulni félreértések. És az is vitathatatlan, hogy ezek megmagyarázása nehézséget fog jelenteni, de ezt a nehézséget vállalni kell azért, hogy hosszú távon eljussunk egy jobban működő, hatékonyabb népegészségügyhöz.

[4] A valóságban ez természetesen nem csak két különböző értéket vehet fel, mint itt a „kis” és a „nagy”, illetve eleve nem csak egyetlen ilyen jellemző van – ahogy az előbbi felsorolás is mutatja – ez most itt a lehető legegyszerűbb példa, hogy megértsük a jelenséget.

Az első amit tegyünk meg, hogy ellenőrizzük le a számokat: nézzük meg, hogy az összegek tényleg stimmelnek (100 + 500 az tényleg 600 stb.), és nézzük meg, hogy az osztások stimmelnek (100 / 5000 tényleg 2% stb.). Ha ezekről meggyőződtünk, akkor láthatjuk: valóban *lehet* ez a helyzet az előző bekezdésben idézett számok mögött. (Fontos a szóhasználat: nem tudhatjuk biztosan, hogy *tényleg* ez van-e a számok mögött, pont ez a lényeg, hogy ha a betegek kockázatát nem ismerjük, akkor ezt nem tudhatjuk. De *előfordulni* előfordulhat ez is, ahogy azt az ellenőrzés igazolta is.)

Mit látunk ezeken az adatokon? Egy első ránézésre teljesen meglepő, sőt, szinte érthetetlen dolgot: az derül ki, hogy a kis kockázatú betegek körében X a jobb, *és* a nagy kockázatúak körében *szintén* X…! De ez mégis hogyan lehet?! – kérdezhetné valaki. Hiszen a betegek – ebben az egyszerű példában – a két csoport valamelyikébe tartoznak, akkor meg hogyan lehet, hogy külön-külön mindkét csoportban jobb X, de összességében *mégis* rosszabb…? Akár azt is mondhatja valaki, hogy ez így egész egyszerűen lehetetlen; mármint nem infektológiailag, hanem matematikailag… De, mint ez a példa mutatja, ez nem igaz, ez igenis lehetséges. De mi erre a magyarázat?

A táblázat rövid tanulmányozása megadja a választ erre a kérdésre. Két dolog tűnhet ugyanis fel:

1. A kockázat *önmagában* is növeli a kórházi fertőzés előfordulását. (Önmagában, értsd: mindegy, hogy a beteg melyik kórházban van.) A 10 nagyobb mint a 2, és 18 nagyobb mint a 3. Ez teljesen logikus és észszerű, épp ezt jelenti a kockázat fogalma.

2. Az, hogy az egyes kórházak milyen betegeket kezelnek, *nagyon* eltérő: X kórház betegeinek *fele* nagy kockázatú, Y-nál viszont alig van ilyen!

És ez a kettő együtt már tökéletesen meg is magyarázza, hogy mi történt: arról van szó, hogy X „gyűjti” a nagy kockázatú betegeket (lehetséges például, hogy ez egy nagy, országos intézmény, ami pont azt jelenti, hogy a komplex beavatkozásokat igénylő, szövődményes, súlyos állapotban lévő betegeket kifejezetten oda kell küldeni), míg az Y egy, úgy szokták mondani: alacsonyabb progresszivitási szinten lévő intézmény, amiben épp azért van kevesebb ilyen, mert ezeket átküldte X-nek. Így végeredményben előáll az a helyzet, hogy X hiába produkál jobb mutatókat mindkét kategóriában, összességében mégis rosszabb lesz az eredménye, mert az összesített statisztikáját „lerontja” a sok nagy kockázatú beteg, akik fertőzési aránya kockázatukból – és nem a kórház tevékenységéből! – fakadóan nagyobb. Hiába produkál az ő körükben *is* jobb arányt, ez mégis rontó hatású lesz, hiszen bármennyire is jól dolgozik X, azért még így is igaz, hogy a nagy kockázatúak fertőződési aránya nagyobb ott, mint az Y-ban a kis kockázatúaké.

Fontos még egyszer kiemelni, hogy a kockázat címszó alatt olyan tényezőkről van szó, amikről a kórház nem „tehet”. Más lenne a helyzet, ha ezek magából a kórház tevékenységéből fakadnának[5], de arra a kórháznak nincsen ráhatása, hogy hány éves betegek

[5]És akkor is más lenne a helyzet, ha nem függnének össze a fertőzések előfordulásával, vagy lenne ugyan összefüggés, de az fordított irányú hatásból származna, de itt nyilván nem ezekről van szó: idősebb korban több a kórházi fertőzés, és valaki nem lesz idősebb attól, mert kórházi fertőzés lépett fel nála.

érkeznek oda. Innentől kezdve viszont, ha az idősebb életkor nagyobb fertőzési kockázatot jelent (mint ahogy tényleg nagyobbat jelent), akkor „méltánytalan” lenne elszámolni a kórház kárára a nagyobb fertőzés előfordulásnak azt a részét, ami a betegeinek idősebb életkorából fakad. (Amint láttuk, ez akár a teljes többlet is lehet, sőt, akár annál is több.)

Fontos, hogy a különbségekből leválasszuk azt a részt, amiről nem tehet a kórház, de a maradékkal is vigyázni kell: amiről a szó statisztikai értelmében tehet a kórház, az nem feltétlenül esik egybe azzal, amiről a szó hétköznapi értelmében tehet, ugyanis az előbbiben benne vannak a kórház által nem, pláne gyorsan, helyi döntésekkel nem befolyásolható tényezők (például a kórházi épület adottságai). Ez azonban nem jelenti azt, hogy ne lenne fontos ez a leválasztás, egyrészt, mert érdekes lehet az is, hogy az épület adottságaihoz hasonló, nem könnyen befolyásolható tényezőknek mi a hatásuk, másrészt, mert bár igaz az, hogy a javítási lehetőségek, beavatkozási pontok megállapításához szükség van erre a finomabb lebontásra, de ennek a finomabb lebontásnak is az az alapja, első lépése, hogy meghatározzuk mi a kórháznak *tulajdonítható* hatás.

2. A jelenség értelmezése

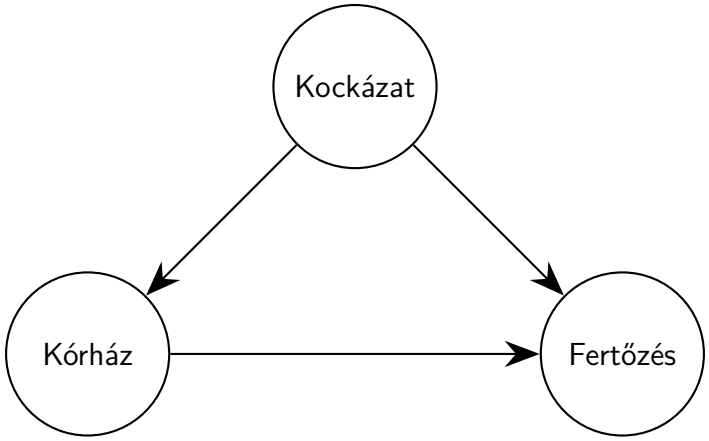
Mit jelent mindez? Azt, hogy az „X kórházban több a kórházi fertőzés” megfogalmazás *rendkívül* félrevezető volt. Nem szó szerint hamis vagy hibás, ezt nem lehet mondani (végülis tényleg több ott a kórházi fertőzés, számszerűen minden adat stimmelt[6]), de mégis, nagyon-nagyon egyértelműen sugallt valamit, ami viszont nem igaz. Ez részben pszichológiai kérdés is: ha a kórház neve mellett tüntetjük fel a fertőződési arányt, akkor automatikusan, óhatatlanul úgy fogja érteni az olvasók többsége, hogy az a kórházat magát jellemzi, és amelyik kórháznál nagyobb – vagy épp kisebb – szám van, ott azt gondolják, hogy az a kórház hatása. Holott, mint láttuk, nem, vagy nem feltétlenül erről van szó: ahelyett, hogy „X kórházban több a fertőzés”, lehet, hogy igazából azt kellett volna mondanunk, hogy „X kórházban *és* nagyobb kockázatú beteg körében több a fertőzés” – ami mindjárt nagyon más!

A probléma lényegében az, hogy a két kórház között nem *csak* az a különbség, hogy az egyik az egyik kórház, a másik meg a másik – eltérnek a kezelt betegek kockázatában is, márpedig innentől kezdve, ha találunk is különbséget a két kórház között, akkor nem tudhatjuk, hogy az mi miatt van: az általunk vizsgált eltérés miatt (melyik kórházban vagyunk), vagy az ezzel *együtt járó* egyéb eltérés vagy eltérések miatt (betegek kockázata), vagy esetleg ezek valamilyen keveréke miatt…? Ez egy általános jelenség az orvostudományban: ha két olyan csoportot hasonlítunk össze, amik nem csak egy tényezőben térnek el, akkor onnantól kezdve, ha találunk is különbséget köztük, nem fogjuk tudni, hogy az melyik eltérés vagy eltérések következménye[7].

[6] Érdekes, de ez bizonyos szempontból talán még rosszabb is, mert egy hamisított adatnál meg lehet mutatni, hogy hamisított, itt viszont több oldalnyi magyarázatot igényel, hogy mi a probléma…

[7] Ezt a jelenséget magyarul is általánosan használt angol szóval confounding-nak szokták hívni. Ez nagyon szemléletes, mert a szó nagyjából olyasmit jelent, hogy „egybemosódás”: az a probléma, hogy az általunk vizsgált eltérés össze van mosódva egyéb eltérésekkel.

A dolgot úgy is megfogalmazhatjuk, hogy van egy *harmadik* változó a történetben: a kórház és a fertőzés mellett ott van a beteg kockázata, ami két dolgot tud, egyszerre hat a kórházra (kockázatosabb beteg valószínűbb, hogy X-be kerül) *és* a kimenetre (kockázatosabb beteg valószínűbb, hogy kórházi fertőzést fog elszenvedni). Az ilyen változót szokás zavaró változónak[8] hívni. Mindezeket egy ilyen ábrával is szemléltethetjük (nagyon szép szóval ezt szokták kauzális diagramnak nevezni):



Amikor tehát arról beszéltünk korábban, hogy „leválasztjuk azt a részt, amiről tehet a kórház", akkor lényegében azt mondtunk, hogy igyekszünk megragadni a valódi *okozati* hatását[9] a kórháznak, megtisztítva a zavaró változó hatásától. Ez az okozati hatás az, amit korábban úgy hívtunk, hogy a kórháznak tulajdonítható hatás, és ami minket érdekel, aminek meghatározása a fő célunk lesz.

3. Útban a megoldáshoz: a rétegzés

Most, hogy értjük a problémát és a jelentőségét, adja magát a kérdés: mit tudunk ez ellen tenni? Mi a megoldás erre?

Kezdjük a jó hírrel: egy megoldást már ismerünk, olyannyira, hogy már használtuk is, és láttuk működés közben – a fenti táblázat a probléma egy lehetséges megoldása! Hiszen, ha soronként nézzük, akkor immár helyes eredményt kapunk (ahogy az történt a korábbi szakaszban is). Ezt a módszert, tehát, amikor a zavaró változó vagy változók összes lehetséges értékére *külön-külön* végezzük el a vizsgálatot, a statisztikusok *rétegzésnek* szokták nevezni (a „réteg" pedig egy ilyen érték, tehát, a táblázat egy sora).

Hogy ez miért működik? Gondoljunk bele, a probléma ugye az volt, hogy amikor összehasonlítottuk a két kórházat, akkor azok eltértek betegösszetételükben *is*. De ez a probléma egy soron *belül* nem jelentkezik, hiszen egy soron belül *minden* beteg kis kockázatú (vagy *minden* beteg nagy kockázatú), így az ilyen, soron belüli összehasonlítását

[8] Magyarul is gyakran használt angol nevén: confounder.

[9] A statisztikai nyelv meg szokta különböztetni az együttjárást (korreláció) és az okozatiságot (kauzalitás). X kórház *együtt jár* a gyakoribb fertőzéssel, de egyáltalán nem biztos, hogy *oka* annak.

a két kórháznak nem torzítja a betegek kockázata, hiszen soron belül mindenki *ugyanaz* a kockázata!

A rétegzés tehát megoldja a problémánkat – legalábbis azon változóra vagy változókra vonatkozóan, amelyek alapján rétegeztünk. Mert ne felejtsük el: a valóságban nem egy zavaró változó van. Ez a valós példában is azonnal látszik, hiszen ha nem a fiktív „kockázat" változót használjuk, akkor rögtön lesz életkor, nem, cukorbetegség, magasvérnyomás-betegség stb. Ezzel kapcsolatban egy dolgot kell észben tartani: ha meg akarjuk őrizni a rétegzés univerzalitását, tehát azon – nagyon kellemes – tulajdonságát, hogy az eredménye helyes lesz, függetlenül a változók viszonyaitól, tehát, hogy semmilyen feltevést nem igényel arra vonatkozóan, hogy mi a változóink közötti kapcsolat jellege, akkor a zavaró változók *összes lehetséges* kombinációjára rétegezni kell. Tehát, ha van cukorbetegség, három értékkel (nincs, 1-es típusú, 2-es típusú) és magasvérnyomás-betegség két értékkel (van, nincs), akkor az $3 \cdot 2 = 6$ lehetséges kombináció (nem cukorbeteg és nincs magasvérnyomása, 1-es típusú cukorbeteg és nincs magasvérnyomása stb.). Vagyis: a számok nem összeadódnak, hanem összes*zorzódnak*! Ha mindemellett van életkorunk, mondjuk 10 lehetséges értékkel[10] (0-9 év, 10-19 év stb.), akkor már $3 \cdot 2 \cdot 10 = 60$ rétegnél tartunk, miközben még csak 3 zavaró változóra rétegeztünk! Mi lesz itt 10 változónál, vagy 20-nál?! Mindenki elképzelheti, hogy milyen számok fognak nagyon hamar kijönni…

Mármost mi történik akkor, ha sok rétegünk, sok sorunk van? Ez egyrészt értelmezési nehézség. Az ember szeretne arra a kérdésre, hogy „mennyivel jobb vagy rosszabb X mint Y" egyetlen számot látni válaszként. Rétegzésnél azonban nem egy válasz van erre a kérdésre, hanem *rétegenként* egy, vagyis annyi válaszunk lesz, ahány rétegünk van. A fenti esetben ez tehát két összehasonlítást jelent, ami nem vészes, de mi van, ha 20 eredményünk van? Vagy 200? Amint láttuk az előbb, egyáltalán nem lehetetlen, hogy ez előforduljon!

A dolog azonban nem csak „kényelmi" problémát jelent, tehát, hogy az értelmezés nehézkesebb, ha sok a réteg. Van egy másik, tisztán statisztikai gond is. A probléma gyökere az, hogy a beteg száma mindeközben adott, állandó (attól még, mert mi több változóra rétegzünk, nem lesz több beteg, az alanyok száma rögzített, a példánkban 13 ezer). Vagyis: ugyanannyi beteg egyre több rétegbe oszlik meg – következésképp egy rétegbe egyre kevesebb beteg fog jutni! Ez azért lesz gond, mert minél kevesebb beteg alapján kell az arányt becsülni, annál bizonytalanabb lesz a kapott eredmény. Hogy a „bizonytalanság" pontosan mit jelent, arra lehetne egészen precíz statisztikai meghatározást is adni, most számunkra bőven elég, ha úgy gondolunk rá, hogy akár egy-két ember státuszának megváltozása is nagyban befolyásolja a végeredményt. Gondoljunk bele, „ezer beteg a tízezerből" és „egy beteg a tízből" egyaránt 10%, de míg az előbbi esetben eggyel több vagy kevesebb beteg szinte semmit nem számít – nem 10% az arány,

[10] Itt nem annyira fontos, de mellesleg ez azt is mutatja, hogy azoknál a változóknál, amiknek nem kategóriák vannak, hanem bármilyen értéket felvehetnek egy tartományon belül, külön probléma van, mert ezek alapján csak úgy lehet rétegezni, ha előbb kategorizáljuk őket, de azt meg nem lehet jól csinálni: ha túl szűk kategóriákat veszünk fel (0-1 év, 1-2 év, stb.), akkor nagyon kevés alany fog egy rétegbe jutni, ha nagyon széleseket (0-40 év, 40-80 év, 80- év), akkor meg egybe fogunk mosni nagyon különböző dolgokat.

hanem 9,99% vagy 10,01% – addig az utóbbinál egyetlen betegnyi változás szó szerint nullára viszi az arányt, vagy épp kétszeresére, 20%-ra emeli.

Tegyük az előbbit egy fokkal azért pontosabbá statisztikailag! Ehhez képzeljük el, hogy van egy cinkelt pénzérménk, amit feldobunk tízezerszer és ezer fejet kapunk. A második esetben feldobjuk tízszer és egy fejet kapunk. Mi a két eset között a különbség? Az ugyanis megegyezik, hogy mindkét esetben azt tippelnénk józan ésszel – ha pusztán ennyi alapján kellene tippelnünk – hogy a pénzérme úgy van cinkelve, hogy 10% valószínűséggel mutat fejet…! Mégis, érezhető, hogy az előbbi esetben ez egy sokkal biztosabb állítás, az utóbbi esetben sokkal bizonytalanabb. De hogyan lehetne ezt megragadni precízen? Egy lehetséges megközelítés, ha a következő kérdésre válaszolunk: „ha a pénzérme valójában 20% valószínűséggel mutat fejet, akkor lehetséges, hogy 10 dobásból 1 fejet kapunk?”. Pici számolással kijön a válasz: igen, könnyedén[11]! (Ugyanis egyáltalán nem biztos, hogy egy ilyen pénz esetén 10-ből *pont* 2 fejet kapunk! Pontosan ugyanúgy, ahogy egy szabályos pénzérmével sem biztos, hogy 10-ből *pont* 5 fejet kapunk; ezt bárki egy pillanat alatt akár ki is próbálhatja.) Menjünk tovább: „és az lehetséges, hogy ha a pénzérme valójában 60% valószínűséggel mutat fejet, akkor 10 dobásból 1 fejet kapunk?”. Újabb rövid számolás után kapjuk az eredményt: ez, ha nem is kizárt, de nem fordulhat elő könnyedén! Ezek tökéletesen megfelelnek a hétköznapi intuíciónknak: 20%-os cinkeltségű pénzérmével könnyen kaphatunk 10-ből 1 fejet, de 60%-os cinkeltségűvel már aligha. És innentől már csak egyetlen lépés van hátra: ne csak két konkrét számra, 10%-ra meg 60%-ra nézzük meg ezt, hanem az *összes* lehetséges valószínűségre 0%-tól 100%-ig! Melyeknél kapjuk azt, hogy olyan valódi cinkeltségnél könnyen előfordulhat az, hogy 10 dobásból 1 fejet kapunk…? A válasz: 0,25%-tól 44,5%-ig. Vagyis, ha a pénzérme cinkeltsége olyan, hogy 0,25%-nál is kisebb a fej-dobás valószínűsége, vagy olyan, hogy 44,5%-nál is nagyobb, akkor nem fordulhat elő egykönnyen, hogy 10-ből 1 fejet dobunk pusztán a véletlen ingadozás miatt. De ha a cinkeltség 0,25% és 44,5% között van, akkor ez könnyedén előfordulhat. A statisztikában ezt a 0,25% – 44,5% intervallumot szokás *konfidenciaintervallumnak* nevezni[12]. Azonnal megérthetjük ennek a jelentőségét, ha kiszámoljuk ugyanezt a másik esetre is, amikor tízezerből kaptunk ezer fejet. Igen, a legjobb tippünk ekkor is 10% – csakhogy ekkor már ennek a konfidenciaintervalluma 9,4% – 10,6%! Nagyobb mintanagyságnál kisebb a véletlen ingadozás, így ez esetben sokkal kisebb eltérések fordulhatnak várhatóan elő. Vagyis: a konfidenciaintervallum szépen visszatükrözi a bizonytalanságot: minél szélesebb, annál bizonytalanabb a kapott eredmény, minél szűkebb, annál pontosabb (legalábbis a véletlen ingadozásra tekintettel).

[11] Ehhez természetesen definiálnunk kell, hogy mit értünk az alatt, hogy „könnyedén”. Most elég annyi, hogy ezt meg lehet tenni statisztikailag szabatos módon; ezt fogja az ún. megbízhatóság jellemezni.

[12] Egész pontosan ez az ún. 95%-os megbízhatósági szinthez tartozó konfidenciaintervallum. Az előző lábjegyzettel összhangban a megbízhatósági szint azt szabályozza, hogy mit értünk az alatt, hogy „könnyedén előfordulhat”. A 95% egy szokásos választás a statisztikában.

4. Egy kitérő: a direkt standardizálás

A probléma klasszikus[13] megoldását a *direkt standardizálás* jelenti.

A módszer alapészrevétele, hogy a „hibás" mutatók (6 és 4%) átlagok – végülis józan ésszel is érezhető, hogy az összes beteg körében mért arány valamilyen formában a kis és nagy kockázatú betegeknél mutatkozó arányok átlaga kell legyen. A csel az, hogy *súlyozott* átlagról van szó: a 4% a 3% és a 18% átlaga, *de* úgy, hogy a kis kockázatú betegek 3%-ának súlya az, hogy milyen arányban vannak kis kockázatú betegek az adott kórházban, hasonlóképp a nagy kockázatúak 18%-ának a súlya az, hogy milyen arányban vannak nagy kockázatú betegek a kórházban. Ellenőrizzük le: $\frac{2800}{3000}\cdot 3\% + \frac{200}{3000}\cdot 18\%$ valóban pont 4%. A dolog teljesen logikus is: amelyik betegből több van, annak nagyobb a szerepe a végeredményben.

És ezen a ponton azonnal láthatóvá válik a probléma: a számok, amiket összeátlagolunk rendben vannak – a gond az, hogy *nem ugyanazzal* a súllyal átlagoljuk őket a két kórházban! Ilyen szemmel nézve látható, hogy épp az hozza létre a problémát, hogy a nagy kockázatú betegek X kórházban 50% súllyal számítanak, Y kórházban viszont kevesebb mint 10% súllyal…!

Mi a megoldás? Ha a probléma az, hogy különböző súlyokkal képezzük a súlyozott átlagot az egyes kórházakban, akkor a megoldás elég kézenfekvően adja magát: használjunk azonos súlyokat! Tehát, induljunk ki a rétegzésből, illetve a rétegzett táblázatban szereplő számokból, hiszen azok rendben vannak, és azt is tartsuk meg, hogy súlyozott átlagot képezünk – csak épp most ügyelve arra, hogy *ugyanazokat* a súlyokat használjuk minden kórház esetében. Ezt az egyezményes súlyozást szokás referenciának vagy standardnak nevezni. Mondjuk, hogy adott a következő referencia-súlyozás: 70% kis kockázatú, 30% nagy kockázatú. Számoljuk ki most a súlyozott átlagokat, ezúttal *mindkét* kórházban *ugyanúgy* ezt a súlyozást használva! Ezzel a módszerrel X kórházban azt kapjuk (úgy szokták hívni: standardizált arányként), hogy $0{,}7\cdot 2\% + 0{,}3\cdot 10\% = 4{,}4\%$, míg Y esetében $0{,}7\cdot 3\% + 0{,}3\cdot 18\% = 7{,}5\%$. Azaz: helyrejött a dolog! Y kórház standardizált mutatója nagyobb, helyesen mutatva, hogy ott rosszabb a helyzet.

Ebben a módszerben az egyedüli kérdés a referencia megválasztása. Ezzel most részletesebben nem foglalkozunk, de annyit azért fontos elmondani, hogy bár lehet érdemi hatása is a végeredményre, de egy dolog biztos: ha egy kórház minden rétegben rosszabb mint egy másik, akkor semmilyen referencia választása mellett nem fordulhat elő, hogy standardizált mutatóban az lesz a jobb.

A szokásos gyakorlatban a direkt standardizálás az elterjedtebb, de egy gyenge pontja mégis van: a bizonytalanság. A probléma már az előző pontban is előkerült, hogy ti. ha kevés beteg van egy rétegben, akkor ott csak nagy bizonytalansággal lehet becsülni az adott rétegbeli arányt (széles lesz a konfidenciaintervallum). Ez a probléma a direkt standardizálással *látszólag* megoldódik, hiszen nem sok számunk lesz, hanem csak egyetlen egy, és a mögé a teljes beteglétszámot be lehet írni, mindenféle szétosztás nélkül. A probléma az, hogy ez csak a látszat: ha valaki utánaszámol, akkor kiderül, hogy a

[13] A direkt standardizálás alkalmazása a 19. század közepére-végére nyúlik vissza. Érdemes megemlíteni, hogy elterjesztésében egy hazánkfia, Kőrösy József statisztikus is komoly szerepet játszott!

valóságban az egyes rétegekbeli bizonytalanság *átöröklődik* a standardizált rátára. Sőt, valójában az a helyzet, hogy nem is csak a rétegek átlagos nagysága vagy hasonló számít – elég *egyetlen egy* nagy bizonytalanságú (kevés betegből számolt) réteg, és már az is nagyon meg tudja növelni a standardizált érték bizonytalanságát[14]. Vagyis ilyenkor, hiába csak egyetlen számról van szó, a standardizált értéknek is nagy lesz a bizonytalansága, széles lesz a konfidenciaintervalluma.

Amint már volt róla szó, klasszikusan a direkt standardizálást szokták alkalmazni első választásként, de ha mégsem, annak jellemzően ez az oka: ha kis méretű rétegek is előfordulnak. A mi konkrét példánkban még ennél is élesebb a helyzet: mint majd látni fogjuk, az új magyar rendszerben olyan finomságú a rétegzés, hogy az is előfordulhat, akár tömegesen is, hogy bizonyos rétegekből nulla beteg volt kórházakban, akár sokban is. Ilyeneknél nem egyszerűen bizonytalan a becslés, hanem lehetetlen (lényegében 0/0 a rétegen belüli arány). Ezért viselte ez a fejezet a „kitérő" címet: a logikai építkezés miatt hasznos, ha az ember ezt a módszert is ismeri, de a jelen esetben nem ez került alkalmazásra.

5. A magyar rendszerben alkalmazott megoldás: az indirekt standardizálás

Az indirekt standardizálást az általános epidemiológiai alkalmazásokban ritkábban használják, mint a direkt módszert, de a kórházi fertőzések vizsgálatában ez a gyakoribb eszköz[15], és a mostani helyzetben is ez lesz a megfelelő választás.

Az alapötlet bizonyos szempontból a direkt standardizálás fordítottja: itt nem a rétegenkénti súlyokra adunk meg egy referenciát, hanem a rétegenkénti *arányokra*. Például ezt mondjuk: kis kockázatúaknál 2,5% a referencia arány a fertőzés előfordulására, nagy kockázatúaknál 11%. Hasonlóan a direkt standardizáláshoz, egyelőre itt se törődjünk azzal, hogy ezek a számok honnan jöttek, vegyük őket adottnak (de, szemben a direkt standardizálással, itt ezt a kérdést később még részletesen meg fogjuk vizsgálni). Mit tudunk ezzel kezdeni? Például ki tudjuk számolni, hogy *ha* egy kórházban minden rétegben a referencia-arány érvényesül*ne*, akkor ott hány fertőzés *lenne*. Az Y kórház esetében például ez így néz ki: 2800 betegünk van a 2,5%-os rétegben, ez várhatóan akkor 70 fertőzött, és 200 betegünk a 11%-osban, ez 22 fertőzött, ami összesen 92 fertőzött. Mit tudunk ezzel kezdeni? Kézenfekvően azt, hogy összevetjük azzal, hogy hány fertőzött volt *tényleg* az adott kórházban! Ha megnézzük a táblázatot, akkor látjuk, hogy Y-ban ténylegesen 120 volt, tehát a helyzet rosszabb, mintha mindenhol (minden rétegben) a referenciának megfelelő eredményt produkált volna. Általában ezt a két érték, a tényleg és a várható hányadosával szokták jellemezni, ez tehát jelen esetben $120/92 = 1{,}3$. Ezt a számot szokták standardizált fertőzési hányadosnak, általánosan használt angol rövidítéssel SIR-nek nevezni; vagy kicsit magyarosabban, a jelen kontex-

[14] Kivéve, ha a rétegnek kicsi a súlya

[15] Ezt a módszert használja többek között az Egyesült Államok és az Európai Betegségmegelőzési és Járványvédelmi Központ is.

tusban hívhatjuk kockázatkiigazított fertőzési hányadosnak is. Minél nagyobb az értéke, annál rosszabb a helyzet a kórházban a referenciához képest. Nézzük most meg ugyanezt X-re is: $5000 \cdot 2{,}5\% + 5000 \cdot 11\% = 675$, vagyis ebben a kórházban a SIR értéke $600/675 = 0{,}89$. Tehát: megint csak helyrejött a dolog, X-ben jobb a helyzet, mint Y-ban! Ez a számítás az *indirekt standardizálás*.

Az indirekt standardizálás hatalmas előnye talán nem annyira ordít a fenti leírásból, pedig benne van: vegyük észre, hogy sehol nem kellett használni a fertőzések *rétegenkénti* arányát! Következésképp mindegy is, ha egy réteg kicsi, illetve, hogy ebből fakadóan az ottani arány nagyon bizonytalan – mert erre az arányra szükség sincs sehol ebben a módszertanban! A fertőzések számából egyedül azok összege kell, még csak azt sem kell tudni, hogy egyáltalán hogyan oszlanak meg a rétegek között. Ebből fakadóan az indirekt standardizálás egy csapásra megoldja a direkt módszer problémáját.

Ettől még természetesen az indirekt standardizálásnál is fontos a bizonytalanság figyelemmel kísérése (hiszen az itt sem nulla lesz, csak kisebb, mint a direkt módszer esetében). Ennek a célszerű eszköze itt is a konfidenciaintervallum – ez ugyanúgy kiszámolható a SIR-re. Például a fenti esetben Y kórház SIR-jének konfidenciaintervalluma 1,08 – 1,56. Ez egy fontos adat, amit mindig a SIR-rel együtt kell értékelni. Ugyanis gondoljunk bele: mi lett volna, ha a konfidenciaintervallum tartalmazza az 1-et? Emlékezzünk vissza a konfidenciaintervallum fogalmára: ez azt jelenti, hogy ha a valódi érték 1 (tehát nincs eltérés, ugyanaz a várható fertőzés-szám mint a tényleges!) *akkor is* könnyedén kijöhetett, pusztán a véletlen ingadozás miatt, az az érték, amit kaptunk. Ez nagyon fontos, hiszen azt mondja: még ha nem is pont 1 a kapott SIR, de mégsincs okunk arra gondolni, hogy itt valódi eltérés van. Ezt szép szóval úgy szokták mondani, hogy nincs *szignifikáns* hatás (a SIR nem tér el szignifikánsan 1-től). Ilyenkor nem állíthatjuk, hogy eltérést találtunk, hiába is nagy vagy kicsi a SIR! A bizonytalanság mértéke a betegek számától függ, tehát simán lehet, hogy valahol az 1,1 is szignifikáns, máshol meg a 2 sem az. Ha azonban a konfidenciaintervallum alsó széle is nagyobb, mint 1 (ahogy a mostani példánkban is van), akkor a kapott eredmény már nem kompatibilis azzal a lehetőséggel, hogy igazából nincs eltérés, és mi pusztán a véletlen ingadozás miatt kaptuk azt az értéket, amit látunk – arra kell következtetnünk, hogy *valódi* eltérés van, a SIR *tényleg* nagyobb mint 1. Fordítva, ha a konfidenciaintervallum felső széle is kisebb mint 1, akkor az érték annyival kisebb, mint 1, hogy azt már nem tudhatjuk be pusztán a véletlen ingadozásnak, arra kell gondolnunk, hogy a SIR a valóságban is kisebb mint 1. Ez utóbbi esetekben mondjuk azt, hogy a tapasztalt eltérés szignifikáns (a SIR szignifikánsan nagyobb vagy kisebb mint 1).

De akkor miért a direkt a gyakrabban használt eszköz? A válasz, nem meglepő módon, az, hogy sajnos az indirekt standardizálásnak van egy nagyon komoly problémája: a számolás során, ahogy a fenti egyszerű példán is látszik, *visszahozza* a kórházak – egymástól különböző! – betegösszetételét, miközben épp ez az, amitől meg akartunk szabadulni. És ennek sajnos következményei vannak. Nézzük meg még egyszer a fenti példát egy apró módosítással: Y kórházban ne 84 eset legyen a kis kockázatú csoportban, hanem csak 60, és ne 36 a nagy kockázatúban, hanem csak 21. Ekkor továbbra is fennáll az alaphelyzet: így is mindkét rétegben Y-ban nagyobb a kockázat, miközben összességében ott kisebb. Mi a helyzet a SIR-rel? A várható érték nem változik (a bete-

gek rétegenkénti számához nem nyúltunk), így a SIR: $(60 + 21)/92 = 0{,}88$. Ami azért nagyon nagy baj, mert – emlékezzünk vissza – az X kórház SIR-je, ami nem változik most, hiszen annak paramétereihez nem nyúltunk, 0,89 volt…! Vagyis, előállt az a helyzet, hogy Y *minden* rétegben nagyobb fertőzési aránnyal bír mint X, de a SIR-je *mégis* kisebb! Tehát előfordulhat olyan helyzet, ahol az indirekt standardizálás *meg sem oldja* azt az alapproblémát, ami miatt egyáltalán az egészbe belekezdtünk. Statisztikailag bebizonyítható, hogy ilyen a direkt standardizálással nem fordulhat elő.

Ez a fajta torzítás[16] sajnos az indirekt módszer elkerülhetetlen velejárója. A jelentősége azonban nem teljesen egyértelmű, hiszen a fentiek „csak" annyit bizonyítottak, hogy ilyen elméletileg előfordulhat (a direkt módszernél ez elméletileg is lehetetlen), de ez semmit nem mond arról, hogy ez mennyire gyakori a valós alkalmazási példákban. A kérdés a szakirodalomban sem tisztázott egyértelműen, de több közlés utal arra, hogy a probléma jelentősége nem túl nagy, és a direkt és indirekt módszer által adott eredmények a legtöbb gyakorlati esetben, ahol mindkettő kiszámolható, valójában nincsenek nagyon messze egymástól.

Az indirekt módszer esetében mindenképp beszélni kell arról a kérdésről, hogy pontosan honnan jön a referencia (emlékezzünk vissza, ennek egy fertőzési arányt kell megadnia minden egyes rétegre). Az egyik lehetőség az *aspiratív* referencia alkalmazása: ilyenkor a referenciát orvosi alapon határozzuk meg, úgy, hogy valamilyen elvárt szintet tükrözzenek a benne szereplő fertőzési arányok; ennek pontos tartalma már ránk van bízva, lehet „minimális", „megfelelő", „ideális" stb. E megközelítés kellemes tulajdonsága, hogy ebben az esetben a SIR maga is egy közvetlenül értelmezhető, abszolút tartalommal bíró mérőszám lesz, hiszen azt fogja megmutatni, hogy az adott kórház hogyan viszonyul az elváráshoz (a minimumhoz, a megfelelőhöz, az ideálishoz stb.).

Az aspiratív referencia problémája, hogy kell valamilyen, lehetőleg megalapozott, védhető orvosi alap, ami indokolja, hogy miért pont annyi az elvárt fertőzési arány az egyes rétegekben amennyi. Ha olyan finomságú a felbontás – márpedig az új magyar rendszerben ez lesz a helyzet – hogy több száz réteg van, akkor ez kivitelezhetetlen lehet, pláne, ha az egyes rétegek tartalma nem bír közvetlen kórházi fertőzésekre vonatkozó relevanciával. Ilyenkor jön szóba a másik módszer, az *empirikus* referencia alkalmazása. Ebben az esetben egyáltalán nem fogalmazunk meg semmilyen elméleti elvárást, hanem megnézzük, hogy a vizsgált kórházakban mennyi a tényadat. Visszatérve az eredeti példánkra: nem elvárást adunk meg arról, hogy mennyi legyen a fertőzési arány a kis kockázatú csoportban, hanem egyszerűen megnézzük, hogy a vizsgált mintánkban (minden kórházban együttvéve!) a 7800 kis kockázatú betegből 184 fertőződött,

[16] Ez a megfogalmazás statisztikailag kicsit pongyola. Valójában arról van szó, hogy ha azt feltételezzük, hogy a két kórház fertőzési arányainak hányadosa minden rétegben ugyanaz – és mi, a konkrét mintánkban, pusztán a véletlen ingadozás miatt látunk eltérő értékeket – akkor ezen hányados becslésére vonatkozóan az indirekt módszernek még kisebb is a torzítása, mint a direktnek. Az indirekt módszer említett problémája akkor jön elő, ha nem ugyanazok az egyes rétegekben az arányok. Ez egy valódi probléma – ahogy a fenti példa is élesen megmutatta – de az igazság az, hogy ha vadul eltérőek az egyes rétegekben a hatások (például kis kockázatúaknál az egyik kórház sokkal jobb, nagy kockázatúaknál a másik), akkor *minden* egy számba sűrítő mutató megkérdőjelezhető lesz, függetlenül az alkalmazott standardizálási módszertől.

így az arány 184/7800 = 2,4%. Hasonlóképp a nagy kockázatú csoportban az arány 536/5200 = 10,3%. Ez az empirikus referencia; ennek az előnye, hogy semmilyen külső, adott esetben vitatható orvosi megfontolást nem igényel. Fontos azonban, hogy ilyenkor *nagyon* más lesz a SIR tartalma: ebben az esetben a SIR már nem egy abszolút mutató, ami egy külső elváráshoz viszonyít, hanem egy relatív mérőszám, ami az *összes* kórház (betegszámmal súlyozott) átlagához hasonlítja az adott intézményt. Ilyen esetben tehát a SIR azt fogja jelenteni, hogy az összes kórház átlagos teljesítményéhez képest hol helyezkedik el a vizsgált kórház. Nagyon fontos, hogy ebből következően a SIR-nek ilyenkor már nincs abszolút tartalma: egyszerűen annyit mond, hogy a összesített átlagához képest hol vagyunk, de hogy az összesített átlag *maga* hol van, borzalmasan magas vagy rendkívül alacsony, arról ez semmit nem mond!

Az empirikus referencián belül meg szokták különböztetni a külső és belső referenciát. Akkor beszélünk belső referenciáról, ha egy adott kórház kiértékelésére használt referencia számításában benne van magának a kérdéses kórháznak is az adata. Ez a legjobban talán akkor érthető meg, ha arra gondolunk, hogy több évnyi adatunk is van; ekkor a belső referencia az, ha a vizsgált év adatait is felhasználjuk, míg a külső referencia az, amikor csak a múltbeli adatokat. A belső referencia használata elméletileg problémás: gondoljunk bele, ha van egy kórház, ami egy bizonyos rétegben nagyon rosszul teljesít, akkor, belső referencia használata esetén, ez a rossz teljesítmény a referenciát magát is meg fogja növelni a kérdéses rétegben[17] – így a kórház teljesítménye mégsem fog annyira kiugróan rossznak tűnni. Úgy is mondhatnánk, hogy a belső referencia 1 felé húzza a SIR-eket.

Az új magyar rendszer indirekt standardizálást használ, empirikus, külső referenciával: a 2023 és 2024 évi adatok alapján számolja a referenciát[18], majd azt használja a 2025 évi adatok indirekt standardizálására. Ez a kezdeti empirikus standard később bővíthető aspiratív elemekkel, például előírható, hogy minden rétegben érjünk el 5% javulást, vagy, hogy bizonyos, kulcsfontosságúnak tekintett rétegekben érjünk el 10%-ot, stb. Ezzel a későbbiekben behozható az aspiratív referencia előnye is a rendszerbe.

6. Az új magyar módszer

Egyetlen kérdést kell már csak tisztázni: mik lesznek a rétegek? Azaz: mik lesznek a zavaró változók, amiket szűrni igyekszünk a kórház valódi hatásának elkülönítése érdekében? (Emlékeztetőül: ezek olyan változók, amikre egyszerre igaz, hogy 1) hatnak arra, hogy a beteg melyik kórházba kerül, *és* 2) hatnak a kórházi fertőzések előfordulására.)

Ahhoz, hogy egy változóra korrigálni tudjunk (bármely módszerrel is), két dolog kell. Az egyik, hogy tudnunk kell róla, hogy az egyáltalán egy zavaró változó, tehát, hogy

[17] Ez különösen igaz akkor, ha egy réteghez csupán kevés kórház ad egyáltalán adatot, vagy az adott réteg gyakoriságát egy kórház nagyon dominálja.

[18] Ez egyúttal mellesleg azt is jelenti, hogy egyáltalán nincs olyan követelmény, hogy a SIR-ek átlaga 1,00 kellene legyen. Simán lehet, hogy mindegyik SIR 1,3, vagy épp 0,8 körüli, de ez egyáltalán nem ellentmondás, sőt, éppen hogy egy hasznos információt ad: ha ez történne, akkor az azt mondja, hogy *globálisan*, tehát minden kórházra kiterjedően romlás – vagy épp javulás – volt 2025-re.

megfelel-e ennek a két kritériumnak. A kérdés nem nyilvánvaló, hiszen csak az előbbit tudjuk empirikusan, a konkrét tényadatok alapján megnézni (az könnyen megvizsgálható, hogy van-e eltérés a kórházak között[19] a betegek életkorában, nemi összetételében, a cukorbetegek arányában stb.), az utóbbi szükségképp orvosi megfontolásokat igényel: az életkor biztos számít, a beteg neme biztos számít, a cukorbetegség minden bizonnyal számít, de vajon a magasvérnyomás-betegség is számít? Számít-e, hogy a betegnek volt korábban infarktusa? Számít a testtömegindexe? A családi kórelőzménye? A probléma az, hogy ezek a kérdések praktikusan a végtelenségig folytathatóak – aligha lesz olyan pont, ahol azt tudjuk mondani, hogy végeztünk, ezek a változók és csak ezek a változók számítanak.

Van azonban még egy probléma. Még ha tudnánk is, hogy mely változók számítanak, még egy dologra szükség van: adattal is kell rendelkeznünk rájuk nézve! Az kevés, hogy elméletileg megállapítjuk, hogy a cukorbetegség számít, csak akkor tudunk rá ténylegesen korrigálni is, ha ismerjük a betegekről, hogy cukorbetegek-e. Fontos: ezt nem csak a kórházi fertőzést elszenvedett betegekről kell tudnunk, hanem az *összesről*…! Hiszen, gondoljunk bele, ha a táblázatunk sorai a „cukorbeteg” és „nem cukorbeteg” kategóriák, akkor minden kórház esetén tudnunk kell, hogy adott kategóriából hányat láttak el *és* közülük hány szenvedett el kórházi fertőzést. Vagyis: minden felvett betegről tudnunk kell a cukorbetegség-státuszát is, azon túl, hogy kapott-e el kórházi fertőzést. Ez nagyon nem mindegy: a kórházi fertőzést elszenvedett betegekről könnyebb részletgazdag információval rendelkezni, mert azokat jelenteni kell a kórházi fertőzéseket gyűjtő rendszerbe. És ez erős limitáció a mai magyar helyzetben: jelenleg semmilyen olyan adatgyűjtés nincs, ami révén az *összes* kórházba felvett betegről tudnánk, hogy ki cukorbeteg. A jobb korrekciós mechanizmusok elérése érdekében fontos továbbfejlesztési lehetőség[20] lenne ezen a helyzeten változtatni.

De mit tudunk tenni addig is? Ezen a ponton érünk el az új magyar rendszer talán leginnovatívabb, világszinten is meglehetősen egyedi pontjához.

Az alapötlet az, hogy használjuk erre a célra a homogén betegségcsoportok (HBCS) rendszerét! De mi az a HBCS…?

A kórházban ellátott betegek mind különbözőek: ha az összes jellemzőjüket tekintjük, nincs két ugyanolyan beteg. Ez azonban kezelhetetlen lenne adminisztratív szempontból, így a ’70-es években elkezdtek az Egyesült Államokban kifejleszteni egy rendszert, melyet 1993-ban Magyarország is bevezetett[21], aminek az a lényege, hogy a betegeket nem túl nagy számú csoport valamelyikébe sorolják, úgy, hogy az egy csoporton belüli betegek a lehető legjobban hasonlítsanak egymáshoz, viszont a különböző csoportba tartozó betegek minél különbözőbbek legyenek egymástól. Ezek lesznek a homogén betegségcsoportok. Például a 001A kódú HBCS a „speciális intracranialis műtétek 18 év felett, nem trauma miatt” kategóriáját jelenti, a 001B a „speciális intracranialis műtétek 18 év felett, trauma miatt”, és így tovább, egészen a 9958-ig, ami a „nyirokrendszer

[19] Márpedig ha van, akkor az irány nem kérdés: a kórházba kerüléstől a beteg nem fiatalodik vagy idősödik meg, tehát a hatás iránya csak az lehet, hogy az életkor hat a kórházra.

[20] Ez nem feltétlenül jelent új adatgyűjtést, annak minden adatszolgáltatói terhével, megfontolható lehetőség a meglevő kórházi jelentések felhasználása erre a célra; vannak jó példák ilyen vizsgálatokra.

[21] Komoly adaptációs munka után, és egyébként az elsők között Európában!

robotsebészeti műtétei"-t jelenti. Jelenleg Magyarországon szűk 800 HBCS létezik (ezek száma változik idővel, néha létrehoznak újat, vagy törölnek régit); 2025-ben összesen 748 HBCS volt, amire kórház legalább egy esetet jelentett. A leggyakoribb HBCS a 0680 („szürkehályogműtét phakoemulsificatios módszerrel, hajlítható műlencse biztosításával"), a legritkább pedig az a 7 HBCS, amibe egyetlen beteget soroltak összesen az országban az egész év alatt, például ilyen a 3800 („felső- és alsó végtagi replantációk, kivéve kéz-, lábujjak").

A HBCS rendszerét a kórházak finanszírozására használják: rendkívül leegyszerűsítve arról van szó, hogy a kórházak az után kapnak pénzt, hogy melyik HBCS-be tartozó betegből hányat láttak el, ugyanis mindegyik HBCS-hez tartozik egy összeg[22]. Fontos azonban, hogy a HBCS-be besorolás nem jelent *pusztán* pénzügyi homogenitást – ha véletlenül ki is derül, hogy egy kislábujj-törés ellátása ugyanannyi pénzbe kerül mint egy fülgyulladásé, akkor sem fognak ugyanabba a HBCS-be kerülni. Az ugyanazon HBCS-be tartozás tehát egyszerre jelent pénzügyi *és* klinikai homogenitást. És ez utóbbit használja ki a kockázatkiigazítás magyar rendszere.

Az ötlet nagyon egyszerű: bár a HBCS-t nem kórházi fertőzési kockázat mérésére találták ki, de van olyan finomságú – a maga több mint 700, klinikailag is nagyjából homogén kategóriájával – hogy arra *is* hasznosítható. Lényegében arról van szó, hogy a képzeletbeli „rizikó" egy – nem tökéletes, de nagyjából is – helyettesítő változójának tekintjük a HBCS-t, azzal az indoklással, hogy a valódi rizikófaktorok nem tökéletesen, de nagyjából leképeződnek a HBCS-ben: akik rizikója kicsi, azok jó eséllyel hasonló HBCS-kbe kerülnek, akik rizikója nagy, azok jó eséllyel más HBCS-kbe. Ez azért kritikus, mert a HBCS-besorolást viszont ismerjük (és, ahogy volt is szó a jelentőségéről, minden betegre!), így ha az előbbi feltételezés helyes, akkor a HBCS-n keresztül meg tudjuk ragadni a rizikót is…! Fontos hangsúlyozni, hogy ez a korrekció nem tökéletes, ilyet nem is lehet állítani, de olyat igen, hogy *valamennyit* várhatóan javít a helyzeten.

Lényegében tehát az fog történni, hogy rétegző változónak a HBCS-t használjuk – egész hasonló táblázatot eredményezve[23], csak épp több mint 700 sorral. Ez a helyette-

[22] Valójában egy ún. súlyszám tartozik hozzá, amiből külön számolással határozzák meg a pénzbeli összeget. Ez azért logikus, mert pénzösszeget rögzíteni értelmetlen lenne, hiszen folyamatosan változik a pénz értéke az infláció miatt, változik az egészségügyre fordítható összeg, így logikusabb azt rögzíteni, hogy az egyes HBCS-knek mi a *relatív* költségességük, tehát, hogy egymáshoz viszonyítva mennyire drága az ellátásuk, mert az sokkal stabilabb jellemző. Jelenleg a „legolcsóbb" HBCS az I. kategóriába sorolt kissebészeti eljárások, 0,05-ös súlyszámmal, a legdrágább HBCS pedig a 999 gramm alatti újszülöttek ellátása 39,2-es súlyszámmal. A rendszer tehát nem mondja meg, hogy konkrétan hány forint egy ilyen kissebészeti beavatkozás vagy egy ilyen extrém koraszülött-ellátás, mert az nagyon változó lenne, de azt megmondja, hogy az utóbbi nagyjából 800-szor drágább mint az előbbi; ami egy jóval stabilabb érték.

[23] Itt most nem annyira fontos, de egy, teljesen technikai nehézség van: az az adatbázis, amibe a HBCS-k jelentése történik, nem kapcsolható össze a kórházi fertőzések adatbázisával, mert nincs közös azonosító a kettőben. Ezen jó lenne a jövőben javítani, addig is, a jelenlegi magyar rendszer a kórház, felvételi dátum, születési dátum, és nem alapján kapcsolja össze a két adatbázist. Ez ugyanis rendelkezésre áll mindkét helyen, és annak a valószínűsége nagyon kicsi, hogy ugyanazon kórház ugyanaznap felvesz két különböző embert, akik véletlenül pont ugyanolyan neműek, és nap pontossággal ugyanakkor születtek. Nagyobb problémát inkább az elgépelések jelentik, amik lehetetlenné teszik az összekapcsolást – kb. 10% azon betegek aránya, akikhez nincsen egyértelmű pár ezzel a

sítő változó: nem az lesz a sorokban, hogy mi a kockázat (ez lenne az ideális), hanem egy olyan dolog, ami szorosan összefügg a kockázattal.

Két óvatosságra intő megjegyzés azonban tartozik mindehhez. Az egyik, hogy azt azért fontos kikötni, hogy a kórházi fertőzés ténye ne hathasson vissza a HBCS-be sorolásra, ellenkező esetben borul az itteni logika. (Ez nem feltétlenül elméleti aggodalom: egyrészt lehet, hogy egy HBCS-nek van olyan változata („szövődménnyel"), amibe a beteg pont a kórházi fertőzés miatt kerül át, de a még fontosabb, ha a kórházi fertőzés maga egy HBCS. Ilyenre példa a véráramfertőzés – a 801A és B jelű HBCS épp ez – ami azt is mutatja, hogy ez a kérdés akár a vizsgált kórházi fertőzéstől is függhet.) A másik potenciális probléma az lehet, ha a kórház a HBCS-be sorolásra is hat. Arról van szó, hogy a zavaró változó olyan, ami hat a kórházra kerülésre, de fordítva nem. Ez vitathatatlanul így van mondjuk az életkor esetében – attól nem lesz valaki idősebb, hogy adott kórházba került, de kerülhet inkább adott kórházba, ha idősebb – viszont a HBCS esetén már nem ennyire vegytiszta a dolog: elképzelhető, hogy a kórház kihat arra, hogy ki milyen HBCS-be kerül. A gyakorlatban ez a hatás vélhetően nem nagy, hiszen a HBCS-t elsősorban mégis csak a beteg alapbetegsége és állapota vezérli, nem a kórház teljesen szabad választása, de a dolgot fontos megemlíteni, mert ha nem teljesül, akkor a HBCS-re történő korrigálással a kórház valódi hatásának egy részét is kiszűrhetjük.

A fenti leírás ugyan abból indult ki, hogy a HBCS használata egy kényszermegoldás („jobb lenne a valódi rizikófaktorok használata, mint például a cukorbetegség, de ezekre sajnos ma Magyarországon nincs adat"), de igazából két hatalmas előnye van a saját jogán is. Az egyik, hogy a HBCS, ha nem is tökéletesen, de valamennyire arról is nyilatkozik, hogy vélhetően milyen beavatkozásokra, műtétekre, eszközhasználatokra kerül sor egy beteg ellátásánál. Ez azért fontos, mert ezek sokszor erős hatással bírnak a kórházi fertőzések előfordulására, miközben a hagyományos rizikófaktorok, mint a cukorbetegség, de akár az alany alapbetegségének ismerete is, nem vagy kevésbé pontosan határozza meg, hogy vajon milyen beavatkozásokra fog sor kerülni. A másik előny az, hogy a HBCS-be történő besorolás *egyébként is* elkészül, vagyis ez minden plusz adatszolgáltatói teher nélkül is rendelkezésre áll! Ez más, HBCS rendszert használó országok számára is releváns szempont lehet, mint egy egyszerű induló-megoldás kockázatkiigazításra (akár csak átmenetileg, amíg nem sikerül kiépíteni a pontosabb, de erőforrás-igényesebb rendszereket).

A HBCS-n kívül mindössze két dolgot tudunk még figyelembe venni (azon kritérium okán, hogy rendelkezésre áll, mégpedig minden betegre): a nemet és az életkort. Ezt az új magyar rendszer meg is teszi; és egyébként ebben megjelenik az indirekt standardizálás egy másik előnye: hogy ezek könnyen beépíthetőek. Megfelelő statisztikai eljárással[24]

módszerrel. Ez jelent tehát némi veszteséget, és emiatt egy kicsit a statisztikai bizonytalanság is nő, de – ha csak nem feltételezzük, hogy az elgépelés összefügg a kórház betegösszetételével – akkor nem torzító.

[24]Az új magyar módszer egy regressziót épít, minden HBCS-re külön-külön, melyben az életkor és a nem a két magyarázó változó. Ezt a regressziót a múltbeli adatokon becsli, majd a tárgyévre ereszti rá, minden betegre kiszámítva egy becsült értéket – ezek kórházankénti összegzésével kapjuk a kórházak várható értékeit. Megengedve az életkor nemlineáris hatását, valamint az interakciót nem és életkor között, egy nagyon flexibilis megközelítést kapunk, ami ráadásul az életkor kategorizálását

megjelennek a várható esetszámban, onnantól pedig a számítás már ugyanaz.

7. Korlátok

Az alkalmazott módszer több statisztikai jellegű limitációval rendelkezik, melyek elsősorban a használt változók struktúrájára, összefüggéseire vonatkoznak. Kiemelendő az, ami egyszerre a módszer erénye és korláta: a homogén betegségcsoportok használata. Ez egyrészt egy erőteljes eszköz (különös tekintettel arra, hogy nem igényel többletadatgyűjtést), de limitációt jelent, hogy mennyire tudja megragadni a beteg-kockázatot. Szintén korlátot jelent a SIR mutatók korlátozott összehasonlíthatósága, noha ennek jelentősége még a szakirodalomban sem egyértelműen tisztázott kérdés.

Az ilyen típusú korlátokat hosszasan lehetne tárgyalni, de ezt javarészt már az előző szakaszok is megtették, emiatt talán célszerűbb itt egy teljesen más jellegű, de egyáltalán nem másodlagos (sőt!) korlátra felhívni a figyelmet: arra, hogy minden tárgyalt módszer, statisztikai eljárás, korrekció *kizárólag* a betegösszetételbeli különbségek kiszűrését célozta, ebből fakadóan semmilyen védelmet nem jelentenek az összes egyéb torzítási forrás ellen. Ezek közül is, kiugró gyakorlati jelentősége miatt, kiemelendő az aluljelentés kérdésköre. Ha egy kórház nem jelenti az eseteit lelkiismeretesen, vagy meg sem találja azokat (bizonyos kórházi fertőzések felismerése nagyon is függhet attól, hogy milyen intenzíven keresik, mennyi mintavételt, tesztet végeznek), az mindenképp félrevezető eredményhez fog vezetni, amin a fenti korrekció semmit nem segít. Ezt a problémát azonban, ha nem is oldja meg, de kontextusba helyezi a módszer alkalmazási célja: amint volt róla szó, az eredmény nem kőbevésett sorrendet jelent, hanem egyfajta orientálását az erőforrásoknak, hogy melyik kórházat érdemes közelebbről megvizsgálni. Ebben a módszer az előbbi limitációjával *együtt is* teljes mértékben hasznos lehet – legfeljebb a vizsgálat eredménye az lesz, hogy a kórház valójában nem kiugróan jól teljesít, hanem kiugróan aluljelenti az eseteit. Ha ez feltárásra kerül, az pontosan ugyanúgy azt jelenti, hogy a módszer működött és betöltötte a célját.

8. Záró gondolatok

Remélhetőleg a fenti tárgyalás is megmutatta, hogy ezek a problémák egyrészt messzemenőkig nem nyilvánvalóak és félreérthetetlenek, másrészt, hogy a megoldásukra szolgáló módszerek nem jelentik aggálytalanul lezárt területeit a tudománynak. A bemutatott módszer maga is bizonyosan továbbfejleszthető, és tovább is fejlesztendő. Emiatt könnyen lehet, hogy mindez vitákhoz is fog vezetni, ami azonban jó hír, mert ezek a viták segítik a további fejlődést. Minél többen vesznek részt ezekben a vitákban, és azok minél jobb színvonalúak, annál valószínűbb, hogy a transzparencia eléri a végső célját: a lakosság egészségi állapotának javítását. Jelen írás ehhez kívánt hozzájárulni.

sem igényli. Ez alól az egyetlen kivételt azok a HBCS-k jelentik, amikben kevés beteg van, és emiatt nem becsülhető ilyen komplex modell; ezeknél a módszer előbb csak életkort használ magyarázó változóként, nemet nem, ha pedig így sem becsülhető, akkor egyszerű átlagot fog használni az adott HBCS-re.

9. Irodalomjegyzék

- S Greenland, H Morgenstern. Confounding in health research. Annu Rev Public Health. 2001(22):189-212. doi: 10.1146/annurev.publhealth.22.1.189.
- AC Wysocki, KM Lawson, M Rhemtulla. Statistical control requires causal justification. Advances in Methods and Practices in Psychological Science, 2022(5):2, doi: 10.1177/25152459221095823.
- TL Lash, TJ VanderWeele, S Haneuse, KJ Rothman (eds.). Modern Epidemiology, 4th edition, Wolters Kluwer, 2021.
- NE Breslow, NE Day. Statistical Methods in Cancer Research. Volume 2. The Design and Analysis of Cohort Studies. International Agency for Research on Cancer. 1987.
- H Inskip, V Beral, P Fraser, J Haskey. Methods for age-adjustment of rates. Stat Med. 1983 Oct-Dec;2(4):455-66. doi:10.1002/sim.4780020404.
- C Rioux, B Grandbastien, P Astagneau. The standardized incidence ratio as a reliable tool for surgical site infection surveillance. Infect Control Hosp Epidemiol. 2006 Aug;27(8):817-24. doi: 10.1086/506420.
- TL Gustafson. Practical risk-adjusted quality control charts for infection control. Am J Infect Control. 2000 Dec;28(6):406-14. doi: 10.1067/mic.2000.109883.
- D Birnbaum, R Zarate, A Marfin. SIR, you've led me astray! Infect Control Hosp Epidemiol. 2011 Mar;32(3):276-82. doi: 10.1086/658331.
- M Delgado-Rodríguez, J Llorca. Caution should be exercised when using the standardized infection ratio. Infect Control Hosp Epidemiol. 2005 Jan;26(1):8-9. doi: 10.1086/503173.
- TL Gustafson. Three uses of the standardized infection ratio (SIR) in infection control. Infect Control Hosp Epidemiol. 2006 Apr;27(4):427-30. doi: 10.1086/503019.
- A Sugár. A standardizálás módszerei és története Magyarországon. Statisztikai Szemle. 2026; 104(6):503-532. doi: 10.20311/stat2026.06.hu0503.
- N Keiding, D Clayton. Standardization and control for confounding in observational studies: a historical perspective. Statistical science (2014): 529-558.
- NE Breslow, NE Day. Indirect standardization and multiplicative models for rates, with reference to the age adjustment of cancer incidence and relative frequency data. J Chronic Dis. 1975 Jun;28(5-6):289-303. doi: 10.1016/0021-9681(75)90010-7.