

A bizonyítékok és a végpontok hierarchiája

Soroncz-Szabó Tamás

„... A bizonyíthatóság fogalma [...] nem lehet tökéletes helyettesítője az igazság fogalmának.”

Alfred Tarski

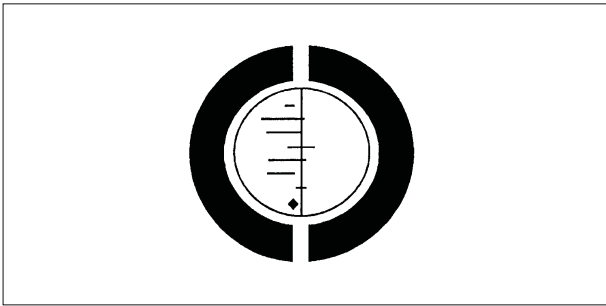
„A z orvoslás legfőbb kötelessége a veszélyeztetett emberi élet megmentése, s a nőgyógyászat pedig az orvoslásnak azon ága, melyben ez a kötelesség igen nyilvánvaló módon teljesítettik” – kezdi Semmelweis Ignác a febris puerperalis kóroktanáról és megelőzéséről írt tanulmányát (1). Ebben számol be arról, hogy a hatalmas forgalmú, bécsi „Ingyenes Szülőotthon” orvosokat képző, 1. részlegében megfigyelt, 10% körüli perinatalis mortalitást a kontrollként szolgáló 2. részlegben tapasztalható 3% körüli szintre sikerült csökkentenie (ahol csak bábákat képeztek, akik nem végeztek boncolást a szülők vizsgálatára előtt). A baktériumok kóroki szerepét évtizedekkel később Joseph Lister fedezte fel. Semmelweis „cadaverosus részecskékről”, illetve „bomló állati-szerves anyagokról” beszél, amelyek a kézről csak klórmésszel távolíthatók el, szappannal nem. Tehát pontos etiológia nélkül is sikerült megtalálnia a hatásos beavatkozást. Mintegy 70%-os relatív kockázatcsökkenés! Az orvostudomány történetének egyik legelső kvázi-randomizált kontrollos vizsgálatával állunk szemben (2). A kortársak Semmelweis kijózanító bizonyítékait döntően elutasító fogadtatásban részesítették. Spaeth bécsi professzor még három évvel a publikáció után is „a csikósostor irtóztató pattogásának” minősítette, a „magyar ugar”-ra célozgatva (3).

Semmelweis munkája (s vele a kézfertőtlenítés is) később csaknem teljesen feledésbe merült. Ha a fenti halálozási adatokat vesszük alapul, az egyszerű és olcsó klórmeszes antiszeptikumok alkalmazásával minden 100 szülőnőre hét felesleges haláleset (három helyett tíz) számolható. Az 1. részleg évi 3–4000-es szülési statisztikájával számolva ez évente 210–280 szülőnő felesleges halála... Felesleges, mert a bizonyítékok mérlegelésével elkerülhető lett volna.

Késlekedő gyakorlat – felesleges szenvedés, halál

A bizonyítottan hatásos beavatkozások gyakorlati alkalmazásának elmulasztása – vagy éppen a nem bizonyított, potenciálisan káros gyakorlat be nem szüntetése – végigkísértik az orvoslás fejlődésének útját mind a mai napig, még a fejlett világban is, a betegek nagy tömegeinek felesleges szenvedést és halált okozva (4–6).

Ennek megelőzésére alakult egy nonprofit nemzetközi szervezet, a Cochrane Együttműködés, amely az egyes klinikai területeken rendelkezésre álló adatokat folyamatosan gyűjti és protokollokban meghatározott, standard módszerek szerint szintetizálja (7). Ennek eredménye a számtalan szisztematikus, vagyis módszeresen elvégzett áttekintés (systematic review). A szisztematikus áttekintés a legeredményesebb módszer a tudományos bizonyítékok felkutatására, kiértékelésére és összesítésére, azzal a feltett szándékkal, hogy a tévedés lehetőségét (a szisztematikus és random hibákat) minimalizáljuk* (a jegyzeteket l. a Zárszó után – *a szerk.*). Ezek egy részében metaanalízis (kvantitatív szintézis) is lehetséges, amelynek során az áttekintett tanulmányok egyedi eredményeit matematikailag egyetlen számadatba összesítik. A Cochrane Collaboration emblémáján a két szembe fordított C által bezárt körben egy ilyen metaanalízis sematizált adatai láthatók (1. ábra). A fenyegető koraszüléskor az anyának adott rövid időtartamú (és olcsó!) kortikoszteroid-kezelés hatását a pretermatikus újszülöttek mortalitására 1972-től tíz év alatt összesen hét placebokontrollos vizsgálat tanulmányozta: ezek eredménye egy-egy vízszintes vonalként ábrázolódik. Minél megbízhatóbb a vizsgálat eredménye, annál rövidebb a vonal. Önmagában egyik vizs-



1. ábra. A Cochrane Collaboration emblémája

gálat alapján sem tehető egyértelmű ajánlás a kezelésre, hiszen minden vízszintes vonal a placebo és a szteroid egyforma hatását jelző függőleges vonal körül szór. Az alsó rombusz azonban a hét vizsgálat összesített eredményét mutatja: nagy biztonsággal állítható, hogy a kortikoszteroid csökkenti a koraszülöttek mortalitását (7, 8). Ha ezt a szisztematikus áttekintést nemcsak 1989-ben végezték volna el (8, 9), hanem 1982-ben, amikor már mind a hét vizsgálat eredménye rendelkezésre állt, akkor koraszülöttek tízezreit tudták volna megmenteni a haláltól és a szükségtelen szenvedéstől, és méregdrága intenzív eljárások költségeit lehetett volna megspórolni (7). Ez a metaanalízis volt az egyik legelső a klinikai gyakorlatra (és az emberéletekre) nagy hatást gyakorló Cochrane-féle szisztematikus áttekintések sorában. A mindenkori bizonyítékok módszeres gyűjtése és összefoglalása – és az újabb adatok integrálásával azok rendszeres frissítése – beláthatatlan hasznára válhat a szenvedő betegeknek (10).

Osteoporosis-irányelvek az infarktusos betegeken okulva

2000. március végén hosszú előkészítés után az amerikai egészségvédelmi és -nevelési intézet (National Institutes of Health, NIH) háromnapos konferenciára hívta össze az osteoporosis témakörének legnagyobb honi szaktekinvéit, hogy az NIH Consensus Development programjának keretében áttekintsék a szakterületen felhalmozódott ismereteket. Két nyílt napon át 32 nagyszerű előadó foglalta össze a betegség epidemiológiáját, kihatásait, kórismézését, valamint megelőzési és kezelési lehetőségeit, amelyekhez a kongresszus látogatói véleményét fűzhetek (11). Mielőtt tehát a harmadik napon bármilyen konszenzusos állásfoglalás (12) megfogalmazásának nekikezdték volna, előbb felsorakoztatták a bizonyítékokat (akár egy bíróság, amely egy betegség ellátási lehetőségeiről ítéletet hozni). Ítélezni pedig nagy felelősséggel jár, főként, ha a populáció jelentős részét érintő beteg-

ség ellátásáról van szó. Nem meglepő tehát, hogy az előadások háttéréül szolgáló bibliográfia 2449 referenciát számlál tematikus csoportosításban (13).

A terápiás lehetőségek áttekintésére megalakult az Osteoporosis Research Advisory Group (ORAG), amely az ottawai McMaster Egyetem statisztikai módszertani csoportja, a Cochrane Musculoskeletal Group, továbbá kilenc olyan szakértő részvételével jött létre, akik különféle osteoporosis-vizsgálatokban részt vettek. Gordon Guyatt, a McMaster Egyetem Klinikai Epidemiológiai és Biostatistikai Tanszékének professzora, a szemléletváltás vezérhajójának, az Evidence-Based Medicine Working Groupnak (14) az elnöke koordinálta a gyógyszeres terápiákról fellelhető bizonyítékok szisztematikus áttekintését. Bevezető előadásában (15) egy különös táblázatot vetített fel – az osteoporosistól meglehetősen távol eső szakterületről –, amelyben párhuzamot vont az akut myocardialis infarctusban (AMI) alkalmazott thrombolysissal szerzett bizonyítékok gyarapodása és az AMI terápiás ajánlásaiban a thrombolysis pozicionálása között, az 50-es évek végétől kezdődően (ekkor végezték az első randomizált kontrollált sztreptokinázvizsgálatot) (16, 17). A szerzők a tanulmányok eredményeit egyenként az előzőekhez adva, egy utólagos, kumulatív metaanalízist végeztek; ebből kiderült, hogy az irányelvek még 1979-ben is csak mint kísérleti eljárást ajánlották a thrombolysist AMI-ban, pedig akkor már közel 20 vizsgálatból több mint 5000 beteg adatai álltak rendelkezésre. Ha akkor elvégezték volna a metaanalízist, egyértelműen igazolni tudták volna ($p < 0,001$) a rendkívül kedvező hatást. A sztreptokinázt ekkor a világon egyedül csak a Német Szövetségi Köztársaságban törzskönyvezték (18). Az első terápiás irányelv, amely a thrombolysis rutinszerű alkalmazását ajánlotta, csak akkor született meg, amikor az utólagos metaanalízis (már több mint 18 000 betegen) igen szignifikáns ($p < 0,00001$) eredményt adott, de ezt a szisztematikus összesítést 1986-ban senki sem végezte el. Ezért a kilencvenes évek elejéig további több tízezer beteget vontak be thrombolysistanulmányokba, amelyekben a kontrollcsoport nem részesült az addig már sokszorosan igazolható életmentő beavatkozásban(!), amit egyesek etikai alapon jogosan kezdtek megkérdőjelezni (19). [Kísérteties hasonlóság a Cochrane-embléma fent leírt történetével. Mindeközben a lidokaint széles körben rutinszerűen alkalmazták myocardialis infarctus után a ritmuszavarok megelőzésére, és annak ellenére szerepelt a terápiás ajánlásokban, hogy hatását nem támasztotta alá bizonyíték, sőt, az eredmények inkább a lidokain enyhén káros voltára utaltak (16). Ugyancsak a nyolcvanas évek végére az I. osztályú antiarrhythmiaszerek virágzásuk tetőfokára értek, és csak az Egyesült Államokban

nagyobb pusztítást vittek végbe, mint a vietnami háború: becsülhetőleg 20–70 ezer, akut myocardialis infarctusos beteg indokolatlanul korai halálát okozták, pusztán azért, mert nem tekintették át a velük végzett vizsgálatok eredményeit, amelyeknek egy részét késve publikálták (6).]

Guyatt előadásában kifejtette, hogy ennek az osteoporosis-konferenciának a legfontosabb célja, hogy az osteoporosis területén ne fordulhasson elő a fenti helyzet, amelyben a szívinfarktusos betegek millióit több mint egy évtizeden át megfosztották a kétségtelenül bizonyított hatású thrombolyticus kezelés lehetőségétől, és ugyanakkor a kérdéses hatású lidokainnal, illetve a káros, I. osztályú antiarrhythmias szerekekkel kezelték őket. S mindezt azért, mert senki nem végezte el a mindenkor hozzáférhető bizonyítékok szisztematikus áttekintését. A kezelési irányelveknek sokkal gyorsabban kellene integrálniuk a felhalmozódó bizonyítékokat, hogy azok mielőbb a mindennapi gyakorlat részévé váljanak (20).

A bajok eredete

Az információs forradalom kora

Darwin szerint a fajok eredete az evolúcióban keresendő. Most azonban egy ennél sokkal érezhetőbb és gyorsabban előrehaladó jelenség határozza meg az emberi nem fejlődését: az információs forradalom. Az evolúció helyett ebben a revolúcióban rejlik a bajok eredete, amelynek folytán sokan úgy érzik, hogy nem képesek több információ befogadására (főként azok feldolgozására). Tény, hogy a világtörténelemben eddig valaha élt összes tudós és kutató 80–90%-a ma is életben van (21). Nem az a gond, hogy életben vannak, hanem az, hogy a közzétett orvostudományi kutatási eredmények száma hozzávetőleg 12–14 évente megduplázódik (22) – becslések szerint évi 2 millió szerzeménnyel gazdagodva (23) –, amiből a legnagyobb hagyományokkal rendelkező, ingyenes elektronikus orvostudományi adatbázis, az amerikai National Library of Medicine által gondozott Medline is „mindössze” évi 400 000-et, vagyis egyötödét képes felszippanítani.

A bizonyítékok rendszerezett felhalmozása és közzététele iránti igény egyidős a felelős orvosi gondolkodás kialakulásával. Magyarországon az egyik első ilyen kezdeményezés az Orvosi Tár című folyóirat létrehozása volt. [A szerkesztést a Magyar Tudós Társaság (az MTA elődje) rendes tagja és a magyar orvosi nyelv legnagyobb megújítója (az „orvos”, a „gyógyít”, a „műtét” stb. szavak kitalálója), Bugát Pál, valamint Pest városának főorvosa, Flór Ferenc végezte.] Az Orvosi Tár egyik 1839. évi szá-

mában a szerkesztők az alábbi elveket fogalmazzák meg: „Szerkesztőségünk szűk köréhez képest iparkodott a külföldi literatúrával és folyóirati halommal (journalisticával) közgyelesen haladni, az első részéről főképp olly munkákat jelentvén, mellyek országunk orvosainak leginkább szükségesek [...]. Amit mi a külföldi folyóiratokból közlünk, az nem merő átirat (de papyro ad chartam) hanem, mennyire csak lehetséges volt, iparkodánk a puszta adatokat (datumokat) elvekre (principiumokra) szorítani, visszavezetni, tehát a testnek lelket adni...” (24).

Nem másról írnak itt, mint az orvostudomány eredményeinek kivonatos, kritikus értékeléséről (critical appraisal – mostanában angol magyarázatot teszünk a zárójelek közé a frissen bevezetett magyar kifejezések után, nem pedig latint, amint az a fenti idézetben is megfigyelhető...). A kivonatos kritikus értékelések a tárgyalt vizsgálat standard formátumú strukturált absztraktján (25) kívül egy értelmező kommentárt is tartalmaznak, amely a vizsgálat által megcélzott problémakör összefoglalóját adja, és abban a vizsgálatból nyert új adatokat jelentőségüknek megfelelően elhelyezi. Ez a bizonyítékokon alapuló orvoslás szekunder publikációs folyóiratainak módszere [Evidence-Based Medicine (26), ACP Journal Club (27), Evidence-Based Mental Health (28)], de az internet számos pontján (29–33), valamint egyes szaklapokban [például Canadian Medical Association Journal (34)] is találunk jól megírt és kommentált – ingyen hozzáférhető – kritikus értékeléseket.

Az összességében „leghatékonyabb” információforrások (a legtömörebben szólva a legmegbízhatóbb adatokról) talán a szisztematikus áttekintések kivonatos kritikus kiértékelései volnának, amelyeknek legszakszerűbb gyűjteményét a brit állami egészségbiztosító, az NHS szisztematikus áttekintésekkel és azok terjesztésével foglalkozó központjának (CRD) honlapján található (35), de a Cochrane-absztraktok is ide sorolhatók (36).

A mérsékelt terápiás hatások kora

A gyógyítás hőskorában egy-egy alapvető felismerés nagy hatású és életmentő gyógyszerek kifejlesztéséhez vezetett. Ilyen nagy hatású terápiák használatához igazolásához elegendő bizonyult néhány egyedi megfigyelés (például a penicillin gyógyítja a fertőzéseket). Azonban, amikor a korunkra jellemző, multifaktoriális etiológiájú krónikus betegségek kimenetelét szeretnénk megváltoztatni, legfeljebb mérsékelt hatásokra (15–50%-os relatív kockázat-csökkenés) számíthatunk (37). Részben azért, mert a patomechanizmusban szerepet játszó különféle faktorok közül egy adott kezeléssel rend-

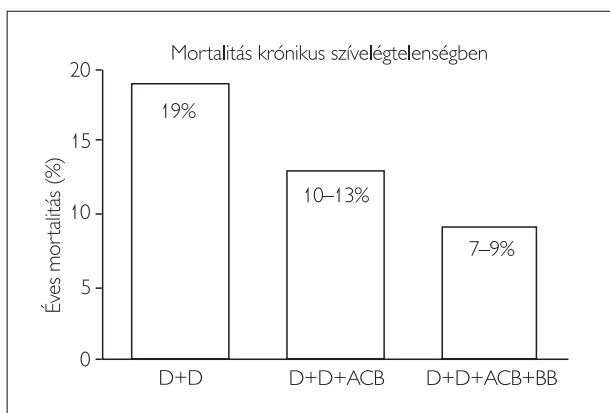
szerint csak egyet tudunk igazán befolyásolni, és a befolyásolható kockázati tényező is (például a vérnyomás) csak mérsékelt összefüggést mutat a lényeges morbiditási kockázatokkal. Így például a 65 évnél nem idősebb hypertoniások között akkor kerülhetjük el 1 stroke felléptét, ha hozzávetőleg 167 betegnek adunk öt éven át folyamatosan antihipertenzív kezelést (38). Egyetlen koszorúér-esemény elkerüléséhez pedig több mint kétezer beteget kell öt évig kezelni (38), hiszen itt még gyengébb az okozati összefüggés. Másrészt azért kell mérsékelt hatással számolnunk, mert ezek a betegségek igen heterogén kimenetellel járnak egy látszólag homogén betegcsoportban is (39).

A kilátások tovább romlanak, ha a megtett első nagy ugrás után tovább akarunk lépni az adott betegség kimenetelének javításában: egyre szerényebb előrehaladásra számíthatunk. A relatív javulás lehet, hogy nem változik az egyes lépések megtétele után [például minden újabb készítmény bevezetésével hasonló nagyságrenddel, újabb 30-40%-kal csökken a krónikus szívelégtelenségben az öszmortalitás kockázata (2. ábra)], de mivel egyre kisebb kiindulási kockázaton érvényesül az a csökkenés, az abszolút (százalékpontos) javulás mértéke egyre szerényebb lesz (ACE-gátló alkalmazásával 6-9 százalékpont, majd a béta-receptor-blokkoló hozzáadásával körülbelül 3-4 százalékpont, vagyis feleakkora a csökkenés) (40-43). Hasonló eredményre jutunk, ha az instabil angina kezelését nézzük (44) (3. ábra). A kisebb abszolút különbség a gyakorlatban azt jelenti, hogy több beteget kell kezelni ahhoz, hogy egy kedvezőtlen kimenetelt megelőzzünk.

Azonban még a mérsékelt terápiás hatás is számíthat az egyes betegeknek, hiszen például melyik instabil anginás beteg ne szeretné, ha a harmincnapos infarktus- és mortalitási kockázata 8%-ról akár csak 5%-ra csökkenne. Sőt, mindez hatalmas ösz-

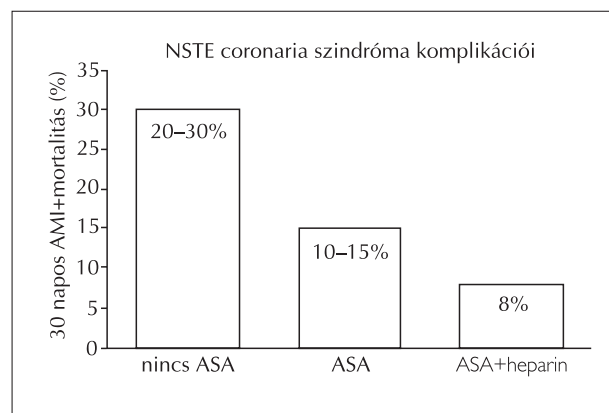
társadalmi haszonnal járhat, ha a kérdéses betegség nagy tömegeket érint. Így a fenti két példában az utolsó állapot szerinti, 8% körüli mortalitási arányok is olyan nagyságrendű betegszámot jelentenek, hogy senki nem kételkedik afelől: van még mit javítani, akármilyen mérsékelt is legyen a hatás! (Csak hazánkban több tízezerre becsülhető az egyik vagy másik kórállapotban szenvedők száma; a 8%-os halálozás alapján közülük évente több ezret veszítünk el.) Szívelégtelenségben a mortalitás további csökkentésére irányuló kísérlet sikertelennek bizonyult egy angiotenzin II-receptor-blokkoló, a valzartán hozzáadásával (ami tüneti javulást okozott a Val-HeFT vizsgálatban, emiatt csökkent a kórházi kezelések gyakorisága) (45). Az ST-eleváció nélküli akut coronaria szindrómás betegek esetén viszont az infarktus és a mortalitás összesített 30 napos kockázata tovább csökkent, ha az eddigi kombinációt kiegészítették tirofibánnal, egy glikoprotein IIb/IIIa-receptor-blokkolóval (PRISM-PLUS vizsgálat) (46). A relatív kockázatcsökkenés hasonlóan bizonyult (40%) az előző terápiák bevezetésével elért kockázatcsökkenésekhez, s ezáltal a 8%-os abszolút kockázat (acetilszalicilsav és heparin) tovább csökkenthető körülbelül 5%-ra. Ez – a fent elmondottaknak megfelelően – kiválóan illeszkedik a 3. ábra oszlopsorába.

A mérsékelt gyógyszerhatások – bár a teljes betegpopulációra éppúgy nagy hatással lehetnek, mint a beteg egyénre – azzal a hátránnyal járnak, hogy meglehetősen nehéz őket kimutatni, létüket igazolni. Mint említettük, a multifaktoriális betegségek prognózisának jelentős egyedi variabilitása miatt megfelelően nagy minta kell, hogy a kimenetel alakulásában kiszűrhető legyen az új beavatkozás okozta tendencia (ha tényleg létezik). Tehát a véletlen hiba esélyét ésszerűen csökkenteni kell. Ugyanakkor számos prognosztikai és egyéb tényezőnek



2. ábra. Balkamra-diszfunkció miatti szívelégtelenség éves mortalitása egyre kisebb mértékben csökkenthető az újabb gyógyszerek hozzáadásával (40-43)

D+D: diuretikum + digitális, ACEI: ACE-gátló, BB: béta-blokkoló



3. ábra. Instabil anginás és ST-eleváció nélküli (NSTEME) infarktusos betegek 30 napos infarktus- és halálozási kockázatának csökkenése újabb kezelések hozzáadásával (44)

AMI: akut myocardialis infarctus, ASA: acetilszalicilsav

akkora a hatása a betegség vizsgált kimenetelére, hogy az összemérhető a tanulmányozott új kezelés várhatóan mérsékelt hatásával. Ha ezek a tényezők bármi ok miatt egyenlőtlenül oszlanak meg az összehasonlított csoportok között, akkor erősíthetik, elfedhetik vagy akár ellenkezőjére változtathatják a vizsgált új beavatkozás által okozott különbséget. Ezeket a szisztematikus hibákat is ki kell küszöbölni.

A mérsékelt hatások korában a klinikai kutatás a random (véletlen) és a szisztematikus (rendszerszerű) hibák elleni küzdeletről szól. A következtetés megbízhatósága attól függ, hogy milyen mértékben sikerült a hibákat kiküszöbölni. Ez a bizonyítékon alapuló orvoslás (BAO; evidence-based medicine, EBM) egyik alappillére.

Behatás és kimenetel: összefüggések a fekete dobozban

A korszerű klinikai orvostudomány alaptudománya az epidemiológia. Epidemiológia alatt már régen nem pusztán járványt kell érteni, és nem is csak a betegségek tér-, idő- és népességbeli megoszlásáról készített kimutatásokat (ez a leíró vagy deskriptív epidemiológia), hanem e megoszlások okait kutató analitikus (elemző), és ezen változtatni szándékozó experimentális (kísérleti vagy intervenciós) tudományt is! Az epidemiológia tehát az empirikus orvostudományi kutatás és oksági bizonyítás módszertana. A gyakorlati betegellátás kapcsán felmerülő diagnosztikus, prognosztikus, profilaktikus és terápiás problémák megoldására alkalmazott epidemiológiát nevezzük klinikai epidemiológiának, mely módszereit tekintve lehet akár deskriptív (pl. a szívinfarktus gyakorisága Magyarországon), analitikus (pl. a szívinfarktus rizikófaktora) vagy experimentális (pl. a szívinfarktus megelőzése gyógyszeres kezeléssel).

Korunk epidemiológiai tudományát a feketedoboz-paradigma jellemzi (47). Magyarul, a klinikai orvostudomány erőfeszítéseinek legjava arra irányul, hogy összefüggéseket keres különböző behatásoknak kitett populációkon a behatások avagy expozíciók (exposures) és a különféle, jó vagy rossz kimenetelek (outcomes) között. A számos hatótényező valamilyen okozati hálón keresztül vezet egy vagy több kimenetelhez. Mindez egy fekete dobozba van zárva, amelynek legjobb esetben csak egy-két „inputját” tudjuk befolyásolni, hogy kedvezően módosítsuk valamelyik kimenetelt. Azt, hogy beavatkozásunknak pontosan meddig terjed a hatása (például: a lipidek módosítása hogyan vezet a koszorúér-események csökkenéséhez; a csontsűrűség emelése mennyiben járul hozzá a törések

1. táblázat. A nem szteroid gyulladáscsökkentők (NSAID-k) szedése attól függően behatás vagy kimenetel, hogy milyen paraméterrel való összefüggését vizsgáljuk

Behatás (expozíció)	Kimenetel
Gyógyszerreklám	NSAID-k szedése
NSAID-k szedése	gyomorvérzés

megelőzéséhez), gyakran jóval később kezdjük megérteni, mint a beavatkozást alkalmazni. Ez a feketedoboz-paradigma lényege (47).

Ha megvizsgálunk egy-két példát az expozícióra és kimenetelre, kiderül, hogy ugyanaz a tényező a vizsgálat céljától függően behatás és kimenetel is lehet (1. táblázat). A kimenetel az egészséggel kapcsolatos esemény vagy állapot (például: makkegészség, zsírfogyasztás, személyiségtípus, cigarettareklám, orvoshoz fordulás, gyógyszereszedés vagy éppen noncompliance, mellékhatás, érlemeszesedés, diabetes mellitus, cataractaműtét, halálozás stb.). A behatás vagy expozíció a kimenetelt befolyásoló bármely tényező (példák: ugyanazok, mint a kimenetelnél).

A citrom útja a gyakorlati alkalmazásig

Vegyünk egy régi esetet! *John Woodall* (William Harvey kortársa) számos megfigyelés alapján már 1639-ben ajánlotta skorbut ellen a citrusfélék fogyasztását (48). A bizonyíték (hogy a citrusfélék megelőzik és gyógyítják a skorbutot) minden bizonnyal azon a megfigyelésen alapult, hogy a citrommal-narancssal szomjukat oltó tengerészek kisebb eséllyel betegedtek meg skorbutban, mint a főként rumot fogyasztó társaik. Később *James Lind* brit orvost a bizonyosság utáni vágy az orvostudomány egyik első kontrollos intervenciós tanulmányának elvégzésére készítette 1747-ben, egy hajó fedélzetén, ahol hat különböző „intervenciós” csoportba sorolt két-két, skorbutban szenvedő matrózt (49). A következő beavatkozásokat alkalmazta: elixir vitriol (fűszerezett kénsav, az akkori standard terápia), almabor, ecet, tengervíz, egy magisztális keverék, illetve napi két narancs és egy citrom fogyasztása (50). Az utóbbi csoportban a matrózok állapota rohamosan javult, míg az öt másik („kontroll-”) csoport matrózainak állapotában nem következett be változás.

Ezzel az experimentális vizsgálattal persze korántsem bizonyosodott be minden kétséget kizáróan a citrusfélék gyógyító hatása, hiszen a csoportonkénti kis betegszám (n=2) miatt az eredmény a tiszta véletlenből is fakadhatott, és a beavatkozásokat sem

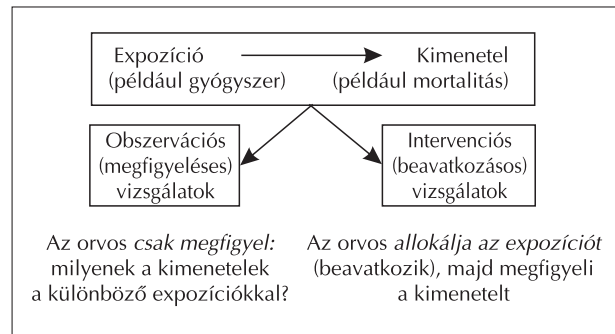
véletlenszerűen (nem random allokációval) osztották ki, így nem lehetünk biztosak abban, hogy a vizsgálat indulásakor egyformán beteg tengerészek voltak-e az egyes csoportokban. (Lehet, hogy a legegészségesebbek kapták a citrusféléket, míg a kénsav, a tengervíz és a többi, kevésbé élvezetes „terápia” a súlyosabbnak látszó betegeknek jutott...)

A véletlen ilyen eshetőségeinek elkerülésével jóval erősebb bizonyítékokat szolgáltatott volna a vizsgálat, de az erre alkalmas módszerek javarészt az orvoslás utóbbi 50 évében fejlődtek ki (37) a klinikai epidemiológia égisze alatt (51), a több okra visszavehető krónikus betegségek előtérbe kerülésével (47). Mindazonáltal az obszervációs és intervenciós bizonyítékok összesítése alapján végül a Brit Tengerészet 1795-ben elrendelte a citrusfélék megelőző célzatú fogyasztását a hajókon (50). [Nota bene: akkor még hat évbe telt, míg Lind publikálta vizsgálatának eredményeit, és közel fél évszázadba, amíg a meggyőzően igazolt terápia bekerült a British Navy „hivatalos ajánlásába”. Manapság csak a negatív eredményű vizsgálatok publikációja szokott évekig elhúzódni (52, 53) vagy teljesen elmaradni (54, 55), de a terápiás ajánlások valamivel hamarabb reagálnak a tudomány előrelépéseire. Ám ez még sajnos korántsem jelenti az igazolt beavatkozás gyakorlati alkalmazásának azonnali elterjedését. Ennek a lassú reakcióidőnek a kihatásairól elég a fenti példákra utalnunk a koraszülöttekkel és a szívinfarktuszos betegekkel kapcsolatban...]

Obszervációs és intervenciós vizsgálatok

A mai klinikai epidemiológiai kutatás célja az expozíció és a kimenetel közötti összefüggések feltárása. Az összefüggések felismerése az orvos egyéni tapasztalatainak eredménye éppúgy lehet, mint a célszerűen az összefüggések feltárására tervezett (klinikai) epidemiológiai tanulmány. Az epidemiológiai vizsgálatok két alaptípusa (56) a vizsgálóorvos és az expozíció viszonyától függ (4. ábra).

1. Ha a vizsgálóorvos passzívan, mintegy kívülállóként elemzi a betegeket ért bizonyos behatások (például a rizikófaktorok vagy az egészségügyi ellátásuk során alkalmazott gyógyszeres vagy más beavatkozások) és bizonyos kimenetek (például a betegek mortalitása) közti összefüggéseket, akkor megfigyelés (obszervációs) tanulmányt végez (amint Woodall tette a citromevés és a skorbut esetében). Ebben a felállásban minden vizsgált személy kezelése adott, amit a beteg és/vagy a kezelő orvos egyenként, „önkéntesen” választott, személyes megítélés és a beteg saját hozzáállása szerint, nem pedig kötelező érvényű hozzárendeléssel, ami



4. ábra. Az epidemiológiai vizsgálatok két alaptípusa a vizsgálóorvos szerepe szerint

az experimentális tanulmányok sajátja. Obszervációs vizsgálatnál sosem lehetünk elég biztosak abban, hogy az összehasonlított csoportok kimenetelei között megfigyelt különbség valóban csak az eltérő expozíciójuknak (például különböző kezelésüknek) köszönhető-e vagy valami más jellemzőben fennálló különbségnek. Az obszervációs csoportok rendszerint az alábbi két alapvető ok miatt nem teljesen egyformák (példákat lásd a következő fejezethez fűzött jegyzetekben):

– Az obszervációs vizsgálatban elkülönített (összehasonlított) csoportok az elkülönítési szemponttal inherens módon összefüggő tulajdonságokkal is rendelkeznek, amelyek a fennálló viszony miatt az összehasonlított csoportok között szisztematikusan eltérnek, így a kimenetelben különbségekhez vezethetnek, de csak akkor, ha önmagukban is prediktorai a vizsgált kimenetelnek. Ezek az úgynevezett zavaró változók (confounding variables). Bár az adatelemzésnél utólag korrigálhatók a csoportok között bizonyos paraméterekben fennálló különbségek – amelyekről úgy gondoljuk, hogy befolyásolhatták a kimenetelt –, abban sohasem lehetünk biztosak, hogy ezek a korrekciók megfelelők, illetve hogy minden olyan paramétert figyelembe vettünk és rögzítettünk, amelyek ténylegesen befolyásolják a kimenetelt.

– A választási lehetőségek „önkéntessége” működni engedi a különféle szelekciós mechanizmusokat, amelyek a betegek saját viselkedése és orvosuk tevékenysége folytán kiválasztódási torzításokhoz (selection bias) vezetnek – vagyis az összehasonlított csoportok nemcsak a vizsgált expozícióban (például az eleve adott kezelésükben) fognak különbözni, hanem ellátásban, odafigyelésben stb.

2. Ha a vizsgálóorvos aktívan részt vesz a behatások (például gyógyszerek) kiosztásában, akkor beavatkozásos (intervenciós) vizsgálatot végez (ahogy Lind kísérletezett a citrom adagolásával). Ez közelíti meg leginkább az „objektív” laboratóriumi kísérleti körülményeket, hiszen itt minden beteg valamilyen besorolási rend szerint kapja a kezelé(s)et, nem pedig szubjektív megítélés alapján. Ezért az in-

tervenció vizsgálat mindig „kísérleti”, vagyis experimentális jellegű. Ha csak egyféle expozíciót vizsgálunk, akkor kontroll nélküli vizsgálatról beszélünk, amelyben a hatás értékelése igen nehézkes, hiszen viszonyítási alapként csak a beavatkozás előtti alapállapot vagy más vizsgálatokból nyert, úgynevezett historikus (történelmi) kontrollok szolgálhatnak. Ez a megközelítés rendszerint a jó obszervációs vizsgálatoknál gyengébb bizonyító erejű adatokat szolgáltat, mert a hibák megfelelő kiküszöbölése lehetetlen (57). Ha azonban legalább kétféle kezelést hasonlítunk össze [placebóval vagy aktív szerrel kezelt csoportokat, attól függően, hogy mi a célunk, és mit tartunk etikus, illetve standard kezelésnek (58, 59)], akkor kontroll vizsgálatról (vagyis kontrollcsoporthoz mint referenciához viszonyító vizsgálatról) van szó. Ha ez a besorolás véletlenszerű (randomizált kontroll vizsgálat), akkor a betegek teljesen egyforma eséllyel kerülnek az egyik vagy másik csoportba, vagyis a tanulmányozott beavatkozásokon kívül minden más körülmény az egyes résztvevők esetében a lehető leghasonlóbb. Ez más módszerrel nem biztosítható két előnnyel jár:

– A tapasztalt hatásokat valóban a vizsgált beavatkozásnak tulajdoníthatjuk és nem olyan betegparamétereknek, amelyek befolyásolják a kimenetelt (zavaró változók). A véletlen besorolással azok a zavaró változók is egyenlően oszlanak meg az összehasonlított csoportok között, amelyekről nem is sejtjük, hogy befolyásolhatják a kimenetelt, vagy amelyeket nem áll módunkban mérni/monitorozni, hogy az elemzésnél a hatásukat korrigálhassuk. Ez obszervációs elrendezésnél lehetetlen.

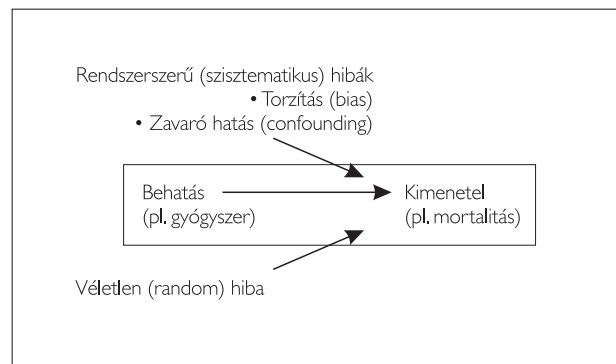
– Kizárható, hogy az összehasonlított beavatkozások hozzárendelését valami egyedi elbírálás vagy a beteg bizonyos jellemzője befolyásolja, ami az összehasonlított csoportok között szisztematikus különbségekhez (torzítás) vezetne.

Statisztikai asszociáció és valós összefüggés

Nem minden statisztikai asszociáció jelent valós összefüggést. (Még akkor sem feltétlenül, ha az összefüggés statisztikailag szignifikáns! Ez a túlmisztifikált „varázsszó” ugyanis csupán azt jelenti, hogy az alább felsorolt négy magyarázat közül a második nem valószínű – de nem is kizárt.)

Ha egy vizsgálatban akár találnak összefüggést, akár az összefüggés hiányát állapítják meg a vizsgált expozíció és kimenetel között, az eredmény minden esetben négyféle módon értelmezhető (5. ábra):

- az eredmény a valóságnak megfelel;
- az eredmény a véletlennek köszönhető (random vagy véletlen hiba);



5. ábra. A kimenetelben bekövetkezett változásnak csak egyik lehetséges oka a vizsgált behatás. Az is lehet, hogy a véletlen és/vagy a szisztematikus hibáknak tulajdonítható a kimenetel alakulása

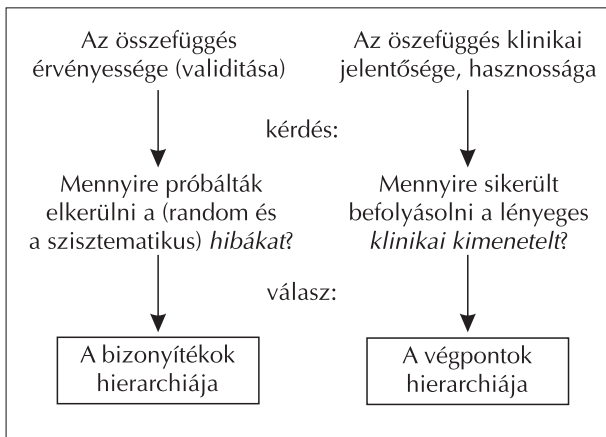
– az eredmény a betegek vagy a vizsgálók által (szándékosan vagy akaratlanul) eltorzított adatok (torzítás: a szisztematikus hibák egyik forrása) miatt jött létre**;

– az eredmény nem a vizsgált behatás és kimenetel közötti viszonyt tükrözi, hanem más zavaró változó(k) (a szisztematikus hibák másik forrása) hatására jött létre, amelyek valamilyen módon együtt járnak a vizsgált expozícióval, és attól függetlenül befolyásolják a kimenetelt***.

Ha tehát egy gyógyszer vizsgálatokor egy kimenetel változását tapasztaljuk (például csökken a cardialis események kockázata), ennek okaként az alkalmazott gyógyszeren kívül szóba jönnek a véletlen vagy a szisztematikus hibafajták. Az eredmények közlésekor mind a négy lehetséges ok szerepét megfelelően érinteni kell. Ahhoz, hogy a vizsgálat validitása maximális legyen (vagyis minél nagyobb eséllyel az elsőként említett eset következzen be), mindent meg kell tenni a további pontokban felsoroltak (random és szisztematikus hibák) kiküszöbölésére. A lehető legteljesebb (de sohasem tökéletes) mértékben ez a randomizált kontrollcsoportos vizsgálatokban valósul meg.

Bizonyítóerő és klinikai relevancia

A bizonyítékokon alapuló orvoslás szemszögéből nézve bizonyítéknak tekinthető minden egészséggel vagy betegséggel összefüggő megfigyelés, amelyet az orvos klinikai gyakorlata során vagy különféle tanulmányok (obszervációs vagy intervenció vizsgálatok) alapján tesz. A terápiás és más irányelvekben olvasható ajánlások erőssége főként az ajánlás alapját képező vizsgálatok felépítésétől (study design) függ, mert a vizsgálatban megfigyelt eredmény bizonyító erejét, „valósághűségét” a vizsgálat



6. ábra. A behatás és kimenetel között megfigyelt összefüggés kritikus értékelése a validitás és a hasznosság megítélésével

felépítése határozza meg (64-68). Másképpen fogalmazva: a vizsgálat felépítésén múlik, hogy a benne megfigyelt összefüggések (például egy rizikófaktor és egy betegség között vagy egy gyógyszer és egy kedvező hatás között) mennyire érvényesek, validak, azaz milyen eséllyel közelíthetjük meg a valóságot. A vizsgálatnak azt az erejét nevezzük validitásnak, amely biztosíthatja, hogy minél valósághűbb eredményt kapjunk. (Egyesek a vizsgálat eredményének általánosíthatóságát – amely egy másik kérdéskör – külső validitásnak nevezik; az itt tárgyalt, szorosabb értelemben vett validitást a vizsgálat belső validitásaként különböztetik el.) A vizsgálatok validitása szerint tehát rangsorolható a belőlük származó eredmények bizonyító ereje, valósághűsége: ezt nevezzük a bizonyítékok hierarchiájának (6. ábra).

Az ajánlásokat tevő testületek azonban két azonos felépítésű vizsgálatból származó, ezért azonos bizonyítóerejű eredmény között is különbséget tesznek. A bizonyítékok hierarchiáján azonos helyet elfoglaló eredmények ugyanis nem feltétlenül azonos klinikai jelentőségűek. Senki nem kételkedne például abban, hogy koszorúérbetege érdekében azokat a gyógyszereket alkalmazza, amelyeknek a legnagyobb bizonyítóerejű vizsgálatokkal igazolódott összmortalitás-csökkentő hatásuk, és ne azokat a gyógyszereket adja, amelyekről ugyanolyan meggyőzően csak annyit igazoltak, hogy csökkennek a plakkok méretét. Azonos erejű bizonyítékok – különböző klinikai relevanciával.

Azt a paramétert, amelyen a beavatkozás (például gyógyszer) hatását a klinikai (gyógyszer) vizsgálat leméri, célváltozónak (end-point variable, elterjedt tükörfordítással végpontnak) nevezik. Ez a gyógyszervizsgálatokban leginkább egy hatásmutató (efficacy measure), vagy egy biztonságossági mutató (safety measure). A tanulmányozott végpontot

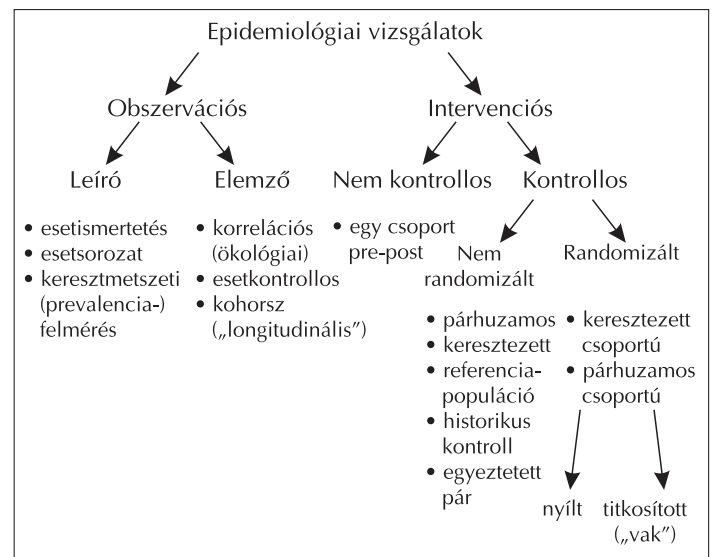
általában akkor nevezik kimenetelnek (outcome – a fent definiált kimenetelnél szűkebb értelemben), ha a betegség lényeges klinikai lefolyását, morbiditási-mortalitási kockázatát mutatja, amelynek komoly egészségügyi kihatásai vannak. Tehát a fenti példát nézve az összmortalitás mint végpont sokkal nagyobb klinikai relevanciájú, mint a plakkregresszió. A tanulmányokból származó eredményeket tehát klinikai jelentőségük, hasznosságuk szerint rangsorolhatjuk: ezt nevezzük a végpontok hierarchiájának (6. ábra).

Bármely orvosi döntéshozatal, s az ahhoz segítő ajánlások alapját tehát e két mérce szerint rangsorolt információk kell, hogy képezzék. Ez a bizonyítékokon alapuló orvosi gyakorlat lényege (69).

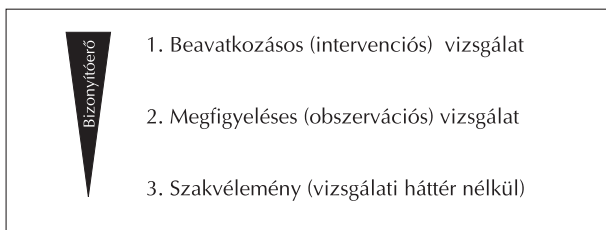
Validabb vizsgálat – valódibb következtetés

A megfigyelés validitásának fő meghatározói: a vizsgálat típusok. Az obszervációs és az intervenciós vizsgálatok fő kategóriáin belül a vizsgálóorvosok több vizsgálati módszer közül választhatnak a rendelkezésre álló idő, erőforrások, illetve a vizsgált probléma természete szerint (7. ábra). A vizsgálatok jellemzőivel, előnyeivel és hátulütőivel az epidemiológiai szakkönyvek kimerítő részletességgel foglalkoznak (56, 64, 70).

Bár szinte minden szakmai szervezet és orvostársaság által megalkotott bizonyítékhierarchia némileg különbözik egymástól, alapjában mindegyik ugyanazt a fő validitási ranglétrát követi (8. ábra) (64-68). A beavatkozással vizsgálatban – amint leírtuk – a vizsgálóorvos kézben tarthatja a



7. ábra. Az epidemiológiai vizsgálatok legfontosabb típusai



8. ábra. A bizonyítékok alapvető hierarchiája eredetük szerint

vizsgált behatás allokációját, így azt kevésbé befolyásolhatják olyan tényezők, amelyek szisztematikus különbséget okozhatnak az exponált és nem exponált csoportok fontos jellemzőiben. Ez utóbbira nagyobb az esély az obszervációs vizsgálatokban, ezért validitásuk kisebb. Az ilyen esetleges különbségek eltorzíthatják a kimenetek alakulását, amelyről így nem lesz pontosan megítélhető, mennyiben függ a tanulmányozott expozíciótól. A vizsgálati háttér nélküli szakvélemény pedig azért szerepel a rangsor legalján, mert alapját az adott kérdést megválaszoló humán klinikai vizsgálatok helyett állatkísérletekből, a patomechanizusból levont következtetésekből, egy másik klinikai kérdést megválaszoló humán vizsgálatokból extrapolált információ képezi, amely – mint minden extrapoláció – bizonyos mennyiségű intuitív összetevőt kell, hogy tartalmazzon ahhoz, hogy az adott kérdésre és betegre formálódjon. Ez jelentősen korlátozza ugyan validitását, de attól még ez is bizonyíték, és gyakran ez a legmagasabb szintű bizonyíték, ami egy adott kérdés megválaszolására rendelkezésre áll!

Az obszervációs vizsgálatok szerepe

A randomizált vizsgálatokkal szemben az obszervációs vizsgálatoknak kétségtelen hátránya, hogy a legkifinomultabb statisztikai korrekciók mellett is megmarad a reziduális torzítás és zavarás esélye. Továbbá a két vizsgálati elrendezés eleve nem ugyanazon típusú kérdések megválaszolására való. Az obszervációs vizsgálatok nélkülözhetetlenek maradnak a rizikófaktorok hatásának vagy a gyógyszerek mellékhatásának becslésében, a vakcinák gyakorlati hatékonyságának monitorozásában, a randomizált kontroll vizsgálatok tervezésében, hipotézisek felállításában stb. (71, 72). Erősségüket mutatja, hogy több, jól elvégzett obszervációs vizsgálatban megfigyelt hatás később randomizált vizsgálatokban is igazolódott (2. táblázat) (71, 73–78). Ugyanakkor jó néhány friss példa hozható gyengébb validitásuk igazolására is, amelyekben a kimutatott

2. táblázat. Az obszervációs vizsgálatokból származó megfigyeléseket a randomizált kontroll vizsgálatok megerősíthetik (71)

Obszervációs vizsgálatokból származó megfigyelés	A megfigyelést megerősítő randomizált vizsgálat
Az acetilszalícilsav megelőzi az első szívinfarktust	Physicians' Health Study (73)
Az első trimeszterben szedett folsav megelőzi a velőcső-záródási rendellenességeket	Czeizel-féle vizsgálat (74) és az MRC Vitamin Study (75)
Az ösztrogének és a SERM-ek fokozzák a vénás thromboembóliás szövődmények kockázatát	HERS (76), NSABP (77), MORE (78)

3. táblázat. Az obszervációs vizsgálatokból származó megfigyeléseket megcáfolhatják a randomizált kontroll vizsgálatok (72)

Obszervációs vizsgálatokból származó megfigyelés	A megfigyelést cáfoló randomizált vizsgálat
A női hormon-pótlás javítja a koszorúér-betegség prognózisát	HERS (79)
A β -karotin megelőzi a tüdőrákot	ATBC (80) és CARET (81)
Az E-vitamin csökkenti a cardiovascularis kockázatot	HOPE (82)

hatások nem bizonyosodnak be a későbbi randomizált vizsgálatokkal (3. táblázat) (72, 79–82).

A randomizált kontroll vizsgálat szerepe

Az egészségügyi beavatkozások valós hatásának legmegbízhatóbb becslésére mindig a randomizált vizsgálat volt és lesz az arany standard – már a szerzői törzskönyvezéséhez is (83). Az alapellátásban az orvos mindennapi gyakorlata során felmerülő kérdések egy alapos taxonómiai kutatás szerint 64 generikus típusba sorolhatók. A 64 közül a leggyakoribb kérdéstípus így hangzik: „Melyik gyógyszert válasszam X kórképre?” (84). Az orvos beállítottságától és lehetőségeitől függ, hogy milyen módon keres rá feleletet (85), de a legvalószínűbb válaszok a legvalidabb tanulmányokból nyerhetők. Ahhoz pedig, hogy a tanulmány validitása maximális legyen (vagyis minél valószínűbb eredményt produkál-

jon), mindent meg kell tenni a random és a szisztematikus hibák kiküszöbölésére.

A randomizált vizsgálat kezdetén a randomizáció, lefolytatása során pedig a kettős vak kezelés és a beteggel való jó együttműködés (kevés elveszett beteg) előzi meg leghatékonyabban a szisztematikus hibákat. A legigényesebb klinikai tanulmányokban ezért mindent megtesznek a kettős vak kezelés megtartásáért. Ennek alapvető eszköze a független Adat- és Mellékhatás-ellenőrző Bizottság (Data and Safety Monitoring Board, DSMB), amely hozzáférhet az időközi elemzések (interim analyses) és bizonyos klinikai betegparaméterek mérési eredményeihez. A FIT vizsgálatban (86) például a beteget kezelő orvosok nem tudtak a csontsűrűség-adatokról, mert azzal könnyen azonosíthaták volna az alendronátot és a placebót szedő betegeket. Ugyanez vonatkozott a 4S vizsgálatban a koleszterinszintre, amellyel a szimvasztatin azonnal elkülöníthető lett volna a placebótól (87). Hogy a betegek mégis biztonságban legyenek, egy előre meghatározott csontsűrűség-csökkenés (a FIT-ben 8% az első évben, vagy 10% két év alatt) vagy összkoleszterinszint elérése (a 4S-ben 9 mmol/l) esetén a DSMB elrendelte adekvát kezelésüket. A kapott kezeléstől függetlenül (ilyenkor a placebós betegek is aktív készítményt kaptak) a vizsgálat végén a betegek adatait az eredetileg kijelölt csoportban elemezték. A kiértékeléskor ugyancsak lényeges, hogy független, a kezelést nem ismerő szakemberek ítélik meg a vizsgált kimenetek (végpontok) bekövetkeztét (Endpoints Committee), amint az a 4S-ben és a FIT-ben egyaránt történt (86, 87).

Az egyedi, randomizált kontroll vizsgálatokkal kapcsolatban megválaszolendő két lefontosabb kérdést a 4. táblázat tartalmazza (64, 69, 88, 89).

4. táblázat. Egy véletlen besorolásos, kontroll vizsgálatlaltal kapcsolatban megválaszolendő két legfontosabb kérdés (és a részletesebb megítélési szempontok) (64, 69, 89)

1. Valóság-hű-e a vizsgálat eredménye?

Ennek legalapvetőbb megítélési szempontjai:

- Megfelelően randomizált volt-e a vizsgálat?
- A betegek követése mennyire volt hiánytalan és megfelelő tartamú?
- Az elemzéskor a betegeket abban a csoportban vették-e számításba, amelyekbe eredetileg sorolták?

Ha az első kérdés igennel megválaszolható:

2. Jelentős-e a vizsgálat eredménye?

Ennek legalapvetőbb megítélési szempontjai:

- Mi a klinikai relevanciája a megfigyelt hatásnak?
- Milyen nagyságrendű a beavatkozás hatása?
- Milyen pontosságú ez a becslés?

A vizsgálat eredménye

Valóság-hű-e? – Validitás

A rosszul kivitelezett randomizált vizsgálatban is következhetnek be hibák. Ezek megítélésére szolgál a 4. táblázat első kérdése. Az első alpontról annyit mindenki sejt, hogy a véletlen besoroláshoz nem megfelelő módszer sem a pénzfeldobás, sem az alternáló beválasztás (például a születési dátum páros vagy páratlan volta szerint). Az orvos és a beteg elől egyaránt eltitkolt (concealment) allokációs módszer kell (90, 91), amelynek mikéntjéről és előnyeiről több helyen olvashatunk (57, 92, 93). Az adekvát randomizáció után a kiindulási csoportok között nem szisztematikus, hanem random különbségek maradhatnak (főként kisméretű randomizált vizsgálatokban, ahol a mintavételi hiba nagyobb lehet). Ettől még igaz a nullhipotézis, miszerint az összehasonlított csoportok ugyanabból a populációból valók (94). Ezek a kiindulási random különbségek (baseline imbalances) azonban statisztikailag könnyen korrigálhatók, így megszüntethető az általa okozott, úgynevezett véletlen torzítás (chance bias). (Ezzel szemben a randomizáció elmaradása- kor jelentkező, betegbeválasztási torzítás által okozott szisztematikus különbségeket utólag korrigálni lehetetlen!) A véletlen kiindulási különbségek statisztikai tesztelése nem informatív (94, 95), ugyanakkor lehetőség nyílik megelőzésükre is, a randomizáció során alkalmazott módszerekkel [rétegzett randomizáció (96) vagy minimalizálás (97, 98)].

A betegek nyomon követését illetően (4. táblázat, 1. kérdés 2. alpontja) figyelembe kell venni egy fontos szempontot. Nem elég az, hogy egy vizsgálatról szóló közlemény a vizsgált betegség prognózisának és szövődményeinek megfelelő hosszúságú követési időt ismerttet: még fontosabb, hogy egyetlen betegnek se vesszen nyoma (loss to follow-up)! Nagy általánosságban, érdemi következtetés nem vonható le olyan vizsgálatokból, amelyek 20%-nál nagyobb követési hiányt szenvedtek (64).

Az egyik randomizált vizsgálatban (99) például az eredményeket szinte érvényteleníti az a tény, hogy az eredetileg bevett betegek 59%-ának nyoma veszett. Ezzel a randomizálás minden előnye is oda-veszett (100). Hiába az adatok elemzése, ha a betegek 59%-áról annyi sem tudható, hogy bekövetkezett-e végponteseménye vagy sem. Az sem jelent semmit, ha mindkét csoportnál ugyanannyi beteg veszett el, mert a két csoport elveszett betegei között valamilyen szisztematikus különbség állhat fenn (akár éppen a vizsgált kimenetek gyakoriságában), amire már sosem derülhet fény. Az elveszett betegeket nem támaszthatja fel semmilyen post factum elemzés, és nem állíthatja helyre a

randomizálás által biztosított egyenlő esélyeket (100). A randomizált vizsgálatban a tökéletlen randomizálás mellett a szem elől veszített (és a rosszul együttműködő) betegek okozhatnak leginkább torzítást.

Bármi, ami a randomizációt követően történik, már befolyásolhatja azt az esélyt, hogy a betegnél bekövetkezzék a vizsgált kimenetel (végpont). Ezért igen fontos, hogy az összes beteg abban a csoportban kerüljön kiértékelésre a vizsgálat végén, amelybe véletlenszerűen besorolták, független attól, hogy szedte-e egyáltalán a vizsgálati szert, sőt, akkor is, ha bármi okból a másik csoport szerét szedte (64). Azok a betegek ugyanis, akik nem felelnek meg a vizsgálati protokoll előírásainak, valamiben szisztematikusan különböznek a protokoll szerint haladóktól. Ha például koszorúérbetegeket nézünk egy randomizált, placebokontrollos vizsgálatban, és az aktív terápiát szedőket jól együttműködők és protokollsértők csoportjára osztjuk, akkor azt fogjuk tapasztalni, hogy az előbbieknél körülbelül feleakkora a halálzási kockázata, és ezt minden bizonnyal a jobb gyógyszeresedésnek tulajdonítanánk. Azon viszont már megdöbbenünk, ha ugyanez volna igaz a jobb placebo-szedőkre a kevésbé szorgalmas placebo-szedőkhöz képest, vagyis amikor egyik alcsoport sem kap hatóanyagot. A CDP vizsgálatban a placebo-ra allokkált betegek között azoknak, akik nem szedték rendesen a placebo-t, körülbelül 1,87-szoros (101), a CAMIAT vizsgálatban 2,25-szoros (102), a BBHAT vizsgálatban pedig 2,5-szeres (103) volt a halálzási kockázatuk (bármely okból) azokhoz képest, akik jó compliance-szal szedték (erősen szignifikáns különbségek)!

A protokollsértő betegek tehát alapjában különböznek a jól együttműködőktől, ezért ezeknek az igen eltérő betegeknek a kihagyása az elemzésből megbillentené a randomizációval kialakult kiindulási egyensúlyt az összehasonlított kezelési csoportok között. Ahhoz tehát, hogy a véletlenszerű besorolás hatása ne vesszen el, az eredeti besorolás szerinti elemzés szükséges (4. táblázat, 1. kérdés 3. pont). Az erre alkalmazott angol kifejezés, az intention-to-treat analysis alatt azonban még a legnevesebb szaklapokban megjelent közlemények szerzői is félrevezető módon mást-mást értenek (104), ezért a módszerek leírásánál kell meggyőződnünk az elemzés mikéntjéről.

Jelentős-e? – A kemény végpontok

A beteg kilátásai és a betegség klinikailag jelentős eseményei (szövődményei) szempontjából a legjelentősebb célváltozókat nevezi az angolszász szakzsargon kemény végpontoknak (hard end-

points, klinikai célváltozók), míg az ezen történéseket több-kevesebb prediktív erővel előre jelző prognosztikus paramétereket pótvégpontoknak (surrogate end-points, indirekt célváltozók) (105). Ezekre néhány alapbetegség kapcsán az 5. táblázatban sorolok fel példákat.

Az orvosok limitált erőforrásainak hatékony kihasználása érdekében *Slawson* és munkatársai azokra a betegorientált végpontokra ajánlják a fókuszálást, amelyek valóban számítanak a betegnek („POEMs”) (89).

A pótvégpontok vizsgálata egyszerűbb, hiszen általában mindenkinél mérhető, nem kell nagy betegcsoportot éveken át addig követni, amíg annyi adat összegyűlik, hogy két kezelés statisztikailag összehasonlítható lesz. Ugyanakkor a pótvégpontokból levont következtetések fatálisan félrevezethetnek (106, 107). Nézzünk egy példát!

Két gyógyszer, amely hasonló betegcsoporton ugyanolyan hatást gyakorol egy pótvégpontra, még nem biztos, hogy a klinikai kimenetelt is hasonlóan befolyásolja. Például mind a bezafibrátról [BECAIT vizsgálat (108)], mind a szimvasztatinról [CIS

5. táblázat. Klinikailag releváns és indirekt célváltozók (kemény és pótvégpontok)

Alapbetegség	Pótvégpont (indirekt célváltozó)	Kemény végpont (klinikai célváltozó)
Atherosclerosis	lumenátmérő, százalékos coronariastenosis	összmortalitás, coronariamortalitás, szívinfarktus
Szívelégtelenség	hemodinamikai paraméterek, terhelhetőség	összmortalitás, hirtelen szívhalál
Osteoporosis	csontsűrűség, csontminőség, a csontforgalom markerei	combnyaktörés, más nem vertebrealis és vertebrealis csonttörések
Hypertonia	vérnyomás	összmortalitás, stroke, szívroham, szívelégtelenség, veseelégtelenség
Asthma bronchiale	FEV ₁	összmortalitás, asthmás rohamok, tüdőszövődmények
Benignus prostata-hyperplasia	prostataméret, reziduális volumen, maximális uroflow, tünetpontszámok	prostatectomia, akut vizeletretenció
Gastrointestinalis gyógyszer-toxicitás	bélpermeabilitás, endoszkóposan detektálható eróziók, dyspepsiás panaszok	gyomor- és nyombélfekély, -vérzés vagy -perforáció, organikus pylorusstenosis

FEV₁: 1 mp alatt erőltetetten kilélegzett levegő

6. táblázat. Két szer azonos hatása a pótvégpontra, de eltérő hatása a kemény végpontokra manifeszt koszorúér-betegségben (108–112)

Végpont	Bezafibrát	Szimvasztatin
„Pótvégpont”		
• Atherosclerosis progressziója	Lassítja (BECAIT)	Lassítja (MAAS, CIS)
„Kemény végpontok”		
• Coronaria-történések	-9,4%, p=0,26 (BIP)	-34%, p<0,0001 (4S)
• Cardialis mortalitás	kb. +7%, p=NS (BIP)	-42%, p=0,0001 (4S)

Zárójelben a klinikai vizsgálatok neve; NS: nem szignifikáns

(109) és MAAS (110) vizsgálatok] kimutatták, hogy koszorúér-betegségben lassítja vagy akár megfordítja a plakkok progresszióját. A klinikai kimenetelre (a kemény végpontra) azonban gyökeresen eltérő módon hatottak a koszorúérbetegeknél [4S (111) és BIP (112) vizsgálatok]. Mindez annak ellenére történt, hogy a bezafibráttal az angiográfiás plakkregresszióra (tehát egy pótvégpontra) kalibrált BECAIT vizsgálatban szignifikáns, 73%-os kockázatcsökkenést tapasztaltak a cardialis eseményekben (6. táblázat).

Számos más területről hozhatunk példákat, miszerint a pótvégpontok – bár alkalmasak lehetnek a betegség stádiumának és progressziójának lemérésére, illetve a rizikófaktorok monitorizálására – gyakran nem képesek előre jelezni a lényeges klinikai kimenetelt (106, 107).

Szisztematikus áttekintések, metaanalízisek

A legmagasabb metodológiai standardok (113, 114) szerint elvégzett, megfelelően interpretált módszeres áttekintésből (és az indokolt esetben elvégzett matematikai adatösszesítésből, azaz metaanalízisből) általában megbízhatóbb és sokrétűbb információk származnak, mint az egyedi tanulmányokból (20). Tény, hogy a jól kivitelezett vizsgálatokat nem pótolhatja a rossz nyersanyagból végzett metaanalízis. Ez ugyanúgy vonatkozik az obszervációs, mint az intervenciós (köztük a randomizált) vizsgálatok metaanalízisére. Az individuális vizsgálatok hibái a metaanalízis eredményében összeadhatnak vagy akár véletlenszerűen ki is olthatják egymást. A szisztematikus áttekintéseknek (köztük a metaanalíziseknek) ezért elengedhetetlen része a vizsgálatok minőségi értékelése (115), különben fennáll a hibás következtetések veszélye (116–119).

Az áttekintett kérdéskörben közölt vizsgálatoknak meg kell felelniük bizonyos követelményeknek ahhoz, hogy a vizsgálatok eredményeit matematikailag szintetizálni lehessen (120, 121). Ha ezek a feltételek nem állnak fenn, akkor a szisztematikus áttekintést célszerűbb metaanalízis nélkül elvégezni. Ekkor ugyanis az egyes vizsgálatok eredményei közötti heterogenitás okainak elemzése, illetve – főként az obszervációs vizsgálatok esetén (122) – a lehetséges zavaró hatások és torzítások elemzése rendszerint több információt ad, mint egy összesített hatásmutató mechanikus kiszámítása (123).

A metaanalízis egyik legfőbb célja az eseményszám növelése a random hiba csökkentése céljából (18, 124). Ugyanakkor a nagyobb eseményszám nagyobb statisztikai erőt eredményez, ami az adott kimenetel tanulmányozásához túlzottan alulméretezett vizsgálatok hamis negatív eredményét billenti helyre (18, 124, 125) (például a Cochrane-embléma vagy a thrombolysis a reinfarctus megelőzésére). A nagyobb mintából eredendően – az eredményektől függően – a metaanalízisből jobban értékelhető az eredmény általánosíthatósága vagy éppen a kisebb betegcsoportokra való érvényessége, mint az egyedi vizsgálatokból (124, 126).

Zárszó

Végletesen fogalmazva, a gyakorló orvos semmiféle vizsgálatból nem tud profitálni, mert azok egy többé-kevésbé homogén populáció viselkedése alapján statisztikai módszerekkel levont következtetéseket adnak, amely vagy túl általános, ha a vizsgálat inkább inkluzív (pragmatikus) felépítésű volt, vagy pedig túl konkrét, ha a vizsgálat inkább exkluzív (válogató) felépítésű. Az inkluzív vizsgálat értelmezése nehéz a konkrét beteg esetére, az exkluzív vizsgálat eredményeiből pedig nehéz a vizsgált betegcsoporton túlmenő, általános következtetéseket levonni (127).

A bizonyítékokon alapuló orvoslás egyáltalán nem váltotta fel a szakvéleményen alapuló orvoslást, hanem két módon kiegészítette azt:

– pontosan helyére tette az epidemiológiai (obszervációs és intervenciós) vizsgálatokon alapuló vélemények, illetve az egyéni tapasztalatokból és az alaptudományokból extrapolált vélemények súlya közötti különbségeket;

– miután az orvost a rendelkezésére álló összes ismeretek rangsorolására ösztönzi az előbbieknél megfelelő szempontok szerint, mindjárt teret enged az egyéni megítélésnek és intuíciónak, hogy a kritikus felmért ismereteket az adott betegre alkalmazhassa – mintegy kitöltve a bizonyítékok hiánya és korlátozott általánosíthatósága miatt létező hatalmas űröket (128). Erre kényszerülnek a legtekin-

télyesebb szakmai fórumok is, amelyek irányelveik megalkotásánál gyakran a korlátozott számú bizonyítékból általánosítva/konkretizálva, vagy trendekből extrapolálva tesznek ajánlást olyan betegalcsoportokra és beavatkozásokra, amelyekre nem áll rendelkezésre meggyőző erejű adat (127). Az ajánlás erejét rendszerint éppen az határozza meg, hogy mennyi mögötte a konkrét bizonyíték, és mennyi az extrapoláció.

A bizonyítékok hierarchiájának tetején az adott kérdést megválaszoló összes releváns randomizált kontrollos vizsgálat jól elvégzett szisztematikus áttekintéséből nyert, megfelelően interpretált következtetés található, mert ez teszi lehetővé a vélet-

len és rendszeres hibák legtökéletesebb eliminációját, illetve azok okainak elemzését. Az újabb és újabb egyedi vizsgálatok eredményeinek integrálása a szisztematikus áttekintés összesített következtetésébe lehetővé teszi a korunkra jellemző, mérsékelt hatású beavatkozások előnyeinek mielőbbi igazolását, s ezzel felesleges szenvedéstől és haláltól menti meg a betegek széles körét. Ugyanakkor feltárhatja az adott eljárás hatástalanságát vagy káros hatását, és annak beszüntetését veheti fel.

A bizonyítékokon alapuló orvosláshoz a klinikai epidemiológia biztosítja a módszertani alapot, amelyet ezért a klinikai orvoslás alaptudományának nevezhetünk (129).

JEGYZETEK

* Az áttekintés és a hiba is szisztematikus? Csak a szó azonos! Az áttekintés esetében a szisztematikus jelző arra a módszerességre utal, amellyel az adatgyűjtés, -értékelés és -szintézis elvégzendő az önkényes és reprodukálhatatlan lépések teljes mellőzésével. A hibánál a szisztematikus szó azt jelzi, hogy a hiba nem véletlenszerűen, hanem rendszeres módon (valamilyen szisztéma szerint) jelenik meg. Tehát: módszeres áttekintés és rendszeres hiba. A „rendszeres” szó időbeli rendszerességre utal („regular”), ezért „szisztematikus” értelemben használata kerülendő.

** Például a menopauza utáni női hormon-pótlás cardiovascularis profilaktikus hatását elemző obszervációs vizsgálatokban (60) torzulást (bias) idézhetet elő, hogy a hormont szedők szorosabb gondozásban részesültek, mint a hormonkezeléssel nem élők – ezért orvoshoz kisebb rendszerességgel járó – társaik, így nagyobb volt a valószínűsége, hogy a rizikófaktorokat és a korai megbetegedéseket felfedezik és kezelik. Ily módon a hormont szedők és nem szedők csoportok szisztematikus eltérő bánásmódban részesültek, ami hozzájárulhatott a hormonszedők csoportjában a

kedvezőbb kimenetelhez (61).

*** Maradva az előbbi példánál, azok a betegparaméterek számíthatnak zavaró változóknak (confounder), amelyek a betegben összefüggnek a hormonszedési statussal (az expozícióval), és ugyanakkor mindentől függetlenül a cardiovascularis (CV) kimenetel prediktorai. Számos CV rizikófaktor tartozik ide, amelynek eredményeként a hormont szedők kiindulási CV statusa eleve szignifikánsan jobb, mint a hormonnal nem élőké. A vizsgálat elején és közben mért zavaró változók (például vérnyomás, obesitas) (62) értékeiben a csoportok között talált különbségek utólag statisztikailag korrigálhatók, de a nem is sejtett vagy más okból meg sem mért zavaró változókból eredő hiba megmarad. (Csak randomizációval érhetek volna el egyforma helyzetet.) Így ad absurdum az is elképzelhető, hogy az alkalmazott kezelés nem hat a cardiovascularis kockázatra, de a zavaró változók kedvezőbb értékei kedvezőbb kimenetelhez vezetnek, mint a nem kezelt csoportban (61–63).

IRODALOM

- Ignaz P. Semmelweis. Die Aetiologie, der Begriff und die Prophylaxis des Kinderbettfiebers. (A gyermekágyi láz etiológiája, fogalma és megelőzése.) Pest-Wien-Leipzig; 1861 (megjelent: 1860).
- Royal College of Physicians of Edinburgh/The Cochrane Collaboration – Controlled Trials from History. http://www.rcpe.ac.uk/controlled_trials/selection.html
- Magyar-Kossa Gy. Orvosi szakoktatás és egyetemi élet a régi Magyarországon. In: M-K. Gy. (szerk.). Magyar orvosi emlékek. Bp. 1929;1:8.
- Mason JO, Koplan JP, Layde PM. The prevention and control of chronic diseases: reducing unnecessary deaths and disability – a conference report. *Public Health Rep* 1987;102(1):17-20.
- Murphy DJ, Povar GJ, Pawlson LG. Setting limits in clinical medicine. *Arch Intern Med* 1994;154:505-12.
- „Registering Information about Randomized Controlled Trials” konferencia. London, 1999. október 4. Sajtóinterjú lain Chalmersszel (UK Cochrane Centre igazgatója). In: Scientists who do not publish trial results are “unethical”. (News) *BMJ* 1999;319:939.
- Cochrane Brochure, 1998. <http://www.cochrane.org/cochrane/cc-broch.htm>
- Crowley P, Chalmers I, Keirse MJ. The effects of corticosteroid administration before preterm delivery: an overview of the evidence from controlled trials. *Br J Obstet Gynaecol* 1990;97(1):11-25.
- Chalmers I, Enkin M, Keirse MJNC (szerk.). Effective care in pregnancy and childbirth. Oxford: Oxford University Press; 1989. p. 759.
- Chalmers I, Haynes B. Systematic Reviews: Reporting, updating, and correcting systematic reviews of the effects of health care. *BMJ* 1994;309:862-5.
- http://odp.od.nih.gov/consensus/cons/111/111_statement.htm
- NIH Consensus Development Panel. Osteoporosis Prevention, Diagnosis, and Therapy. *JAMA* 2001;285(6):785-95.
- <http://www.nlm.nih.gov/pubs/cbm/osteoporosis.html>
- Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA* 1992;268(17):2420-25.
- Guyatt GH. Systematic reviews of osteoporosis therapies. NIH consensus Development Conference on Osteoporosis Prevention, Diagnosis, and Therapy, March 27-29, 2000. National Institutes of Health, Bethesda, Maryland.
- Antman EM, Lau J, Kupelnick B, Mosteller F, et al. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. *JAMA* 1992;268(2):240-8.
- Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, et al. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 1992;327(4):248-54.
- Egger M, Smith GD. Meta-analysis: Potentials and promise. *BMJ* 1997;315:1371-4.
- Geraci E. Unethical placebo assignment in clinical trials of thrombolysis (letter; comment). *J Am Coll Cardiol* 1992;20(5):1302-3.
- Sheldon TA, Guyatt GH, Haines A. Getting research findings into practice: When to act on the evidence. *BMJ* 1998;317:139-42.
- Arndt KA. Information excess in medicine. *Arch Dermatol* 1992;128:1249-56.

22. *de Solla Price DJ*. Little science, big science. *New York: Columbia University Press; 1963. p. 39.*
23. *Lundberg GD*. Perspective from the editor of JAMA, the Journal of the American Medical Association. *Bull Med Library Assoc 1992;80:110-14.*
24. *Bugát Pál, Flór Ferenc*. Orvosi tárunk harmadik félévének megnyitása. *Orvosi Tár 1839;3(1):1-3.*
25. *Begg C, et al*. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA 1996;276(8):637-9.*
26. <http://www.acponline.org/journals/ebm/ebmmenu.htm>
27. <http://www.acponline.org/journals/acpj/cjcmenu.htm>
28. <http://www.ebmentalhealth.com/>
29. Journal Club on the Web. <http://www.journalclub.org/>
30. Family Medicine Research Reviews, Toronto Western Hospital. <http://dfcm19.med.utoronto.ca/twhdfcm/evans.htm>
31. Pediatrics Critical Care Medicine Evidence-Based Journal Club, Washington University. http://PedsCCM.wustl.edu/EBJournal_Club.html
32. Pediatric Evidence Based Medicine, University of Washington. <http://depts.washington.edu/pedebm/>
33. SORAHSN, McMaster University, Ontario, Canada. <http://ahsn.lhsc.on.ca/cat/>
34. <http://www.cma.ca/cmaj/index.htm> (l. "Clinical Update" a News rovatban)
35. <http://nhscrd.york.ac.uk/welcome.htm>
36. <http://www.update-software.com/cochrane/cochrane-frame.html>
37. *Peto R, Baigent C*. Trials: the next 50 years. Large scale randomised evidence of moderate benefits. *BMJ 1998;317:1170-71.*
38. *Gueyffier F, Froment A, Gouton M*. New meta-analysis of treatment trials of hypertension: improving the estimate of therapeutic benefit. *J Hum Hypertens 1996;10(1):1-8.*
39. *Collins R, MacMahon S*. Reliable assessment of the effects of treatment on mortality and major morbidity. I: clinical trials. *Lancet 2001;357(9253):373-80.*
40. *Cohn JN, Archibald DG, Ziesche S, Franciosa JA, et al*. Effect of vasodilator therapy on mortality in chronic congestive heart failure. Results of a Veterans Administration Cooperative Study. *N Engl J Med 1986;314(24):1547-52.*
41. CIBIS-II Investigators and Committees. The Cardiac Insufficiency Bisoprolol Study II (CIBIS-II): a randomised trial. *Lancet 1999;353(9146):9-13.*
42. MERIT-HF Study Group. Effect of metoprolol CR/XL in chronic heart failure: Metoprolol CR/XL Randomised Intervention Trial in Congestive Heart Failure (MERIT-HF). *Lancet 1999;353(9169):2001-7.*
43. *Pitt B, Poole-Wilson PA, Segal R, Martinez FA, et al*. Effect of losartan compared with captopril on mortality in patients with symptomatic heart failure: randomised trial – the Losartan Heart Failure Survival Study ELITE II. *Lancet 2000;355(9215):1582-7.*
44. *Theroux P, Cairns JA*. Unstable angina. In: Yusuf S, Cairns JA, et al (szerk.). Evidence based cardiology. London: *BMJ Publishing Group; 1998. p. 395-416.*
45. *Coats AJS*. The effects of angiotensin receptor antagonists on mortality and morbidity in heart failure – and an interaction with beta blockade. *Int J Cardiol 2001;77:1-4.*
46. The Platelet Receptor Inhibition in Ischemic Syndrome Management in Patients Limited by Unstable Signs and Symptoms (PRISM-PLUS) Study Investigators. Inhibition of the platelet glycoprotein IIb/IIIa receptor with tirofiban in unstable angina and non-Q-wave myocardial infarction. *N Engl J Med 1998;338(21):1488-97.*
47. *Susser M, Susser E*. Choosing a future for epidemiology: I. Eras and paradigms. *Am J Public Health 1996;86:668-73.*
48. *Woodall J*. The surgeon's mate, or military & domestic surgery... London: Printed by Rob Young, for Nicholas Bourne; 1639.
49. *Lind J*. A Treatise of the Scurvy. *Edinburgh: Sands, Murray & Cochran; 1753.*
50. *Thomas DP*. Sailors, scurvy and science. *J Royal Soc Med 1997;90:50-54.*
51. *Vandenbroucke JP*. Evidence-based medicine and "médecine d'observation". *J Clin Epidemiol 1996;49(12):1335-8.*
52. *Stern JM, Simes RJ*. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *BMJ 1997;315:640-45.*
53. *Furberg CD, Psaty BM*. JNC VI guidelines. *Lancet 1998;351(9098):288-93. (in Letters)*
54. *Easterbrook PJ, Berlin JA, Gopalan R, et al*. Publication bias in clinical research. *Lancet 1991;337:867-72.*
55. *Chalmers I*. Underreporting research is scientific misconduct. *JAMA 1990;263:1405-8.*
56. *Hennekens CH, Buring JE*. Epidemiology in Medicine. Ed: Mayrent SL. Boston: Little Brown; 1987.
57. *Altman DG, Bland JM*. Statistics notes: Treatment allocation in controlled trials: why randomise? *BMJ 1999;318:1209.*
58. *Tramèr MR, Reynolds DJM, Moore RA, et al*. When placebo controlled trials are essential and equivalence trials are inadequate. *BMJ 1998;317:875-80.*
59. *Temple R, Ellenberg SS*. Placebo-controlled trials and active-control trials in the evaluation of new treatments. Part 1: ethical and scientific issues. *Ann Intern Med 2000;133(6):455-63.*
60. *Barrett-Connor E, Grady D*. Hormone replacement therapy, heart disease, and other considerations. *Annu Rev Public Health 1998;19:55-72.*
61. *Rossouw JE*. Debate: The potential role of estrogen in the prevention of heart disease in women after menopause. *Curr Control Trials Cardiovasc Med 2000;1:135-8.*
62. *Rödström K, et al*. Pre-existing risk factor profiles in users and non-users of hormone replacement therapy: prospective cohort study in Gothenburg, Sweden. *BMJ 1999;319:890-93.*
63. *Khaw KT*. Hormone replacement therapy again. Risk-benefit relation differs between populations and individuals. *BMJ 1998;316:1842-4.*
64. *Sackett DL, Straus SE, Richardson WS, et al*. Evidence-based medicine. How to practice and teach EBM. 2nd Ed. London: *Churchill Livingstone; 2000. p. 173-7.*
65. *Guyatt GH, Haynes RB, Jaeschke RZ, et al* for the Evidence-Based Medicine Working Group. Users' guides to the medical literature XXV. Evidence-based medicine: principles for applying the Users' Guides to patient care. *JAMA 2000;284:1290-96.*
66. *McKinnell I, Elliott J, Frankish R*. The Cochrane Library. Self training guide and notes. 2000/4. In: The Cochrane Library, Issue 4, 2000. *Oxford: Update Software.*
67. *Eccles M, Freemantle N, Mason J* for the North of England ACE-inhibitor Guideline Development Group. North of England evidence based development project: guideline for angiotensin converting enzyme inhibitors in primary care management of adults with symptomatic heart failure. *BMJ 1998;316:1369-75.*
68. *Gibbons RJ, Chatterjee K, Daley J, et al*. ACC/AHA/ACP-ASIM guidelines for the management of patients with chronic stable angina: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee on the Management of Patients With Chronic Stable Angina). *J Am Coll Cardiol 1999;33:2092-197.*
69. *Straus SE, Sackett DL*. Getting research findings into practice. Using research findings in clinical practice. *BMJ 1998;317:339-42.*
70. *Gordis L*. Epidemiology. *Philadelphia: WB Saunders; 1996.*
71. *Jick H, Garcia Rodriguez LA, Perez-Gutthann S*. Principles of epidemiological research on adverse and beneficial drug effects. *Lancet 1998;352(9142):1767-70.*
72. *Pocock SJ, Elbourne DR*. Randomized trials or observational tribulations? *N Engl J Med 2000;342:1907-10.*
73. Steering Committee of the Physicians' Health Study Research Group. Final report on the aspirin component of the ongoing Physicians' Health Study. *N Engl J Med 1989;321:129-35.*
74. *Czeizel AE, Dudás I*. Prevention of the first occurrence of neural-tube defects by periconceptional vitamin supplementation. *N Engl J Med 1992;327:1832-5.*
75. MRC Vitamin Study Research Group. Prevention of neural tube defects: results of the Medical Research Council Vitamin Study. *Lancet 1991;338:131-7.*
76. *Grady D, Hulley SB, Furberg C*. Venous thromboembolic events associated with hormone replacement therapy. *JAMA 1997;278:477.*
77. *Fisher B, Costantino JP, Wickerham DL, et al*. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *J Natl Cancer Inst 1998;90(18):1371-88.*
78. *Ettinger B, Black DM, Mitlak BH, et al*. Reduction of vertebral fracture risk in postmenopausal women with osteoporosis treated with raloxifene: results from a 3-year randomized clinical trial. [published erratum appears in JAMA 1999;282(22):2124.] *JAMA 1999;282(7):637-45.*
79. *Hulley S, Grady D, Bush T, et al*, for the Heart and Estrogen/Progestin Replacement Study (HERS) Research Group. Randomised trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women. *JAMA 1998;280:605-13.*

80. The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study Group. The effect of vitamin E and beta carotene on the incidence of lung cancer and other cancers in male smokers. *N Engl J Med* 1994;330:1029-35.
81. Omenn GS, Goodman GE, Thornquist MD, et al. Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease. *N Engl J Med* 1996;334:1150-55.
82. The Heart Outcomes Prevention Evaluation Study Investigators. Vitamin E supplementation and cardiovascular events in high-risk patients. *N Engl J Med* 2000;342:154-60.
83. McAlister FA, Laupacis A, Wells GA, Sackett DL, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, XIX: applying clinical trial results, B: guidelines for determining whether a drug is exerting (more than) a class effect. *JAMA* 1999;282:1371-7.
84. Ely JW, Osherooff JA, Gorman PN, et al. A taxonomy of generic clinical questions: classification study. *BMJ* 2000;321:429-32.
85. Soroncz-Szabó T. A bizonyítékokon alapuló orvoslásról, II. rész. Az orvosi döntéshozatal háttere és a hagyományos információforrások. *TAO* 1999;1(4):5-12.
86. Black DM, Reiss TF, Nevitt MC, et al. Design of the Fracture Intervention Trial. *Osteoporos Int* 1993;3(Suppl 3):S29-39.
87. Scandinavian Simvastatin Survival Study Group. Design and baseline results of the Scandinavian Simvastatin Survival Study of patients with stable angina and/or previous myocardial infarction. *Am J Cardiol* 1993;71(5):393-400.
88. Tonkin AM. Evaluation of large scale clinical trials and their application to usual practice. *Heart* 2000;84(6):679-84.
89. Slawson DC, Shaughnessy AF, Bennett JH. Becoming a medical information master: feeling good about not knowing everything. *J Fam Pract* 1994;38(5):505-13.
90. Torgerson DJ, Roberts C. Understanding controlled trials: Randomisation methods: concealment. *BMJ* 1999;319:375-6.
91. Schulz KF. Assessing allocation concealment and blinding in randomised controlled trials: why bother? *Evid Based Ment Health* 2000;3:4-5.
92. Altman DG, Bland JM. Statistics notes: How to randomise. *BMJ* 1999;319:703-4.
93. Schulz KF. Randomized trials, human nature, and reporting guidelines. *Lancet* 1996;348:596-8.
94. Roberts C, Torgerson DJ. Understanding controlled trials: Baseline imbalance in randomised controlled trials. *BMJ* 1999;319:185.
95. Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet* 2000;355(9209):1064-9.
96. Kernan WN, Viscoli CM, Makuch RW, et al. Stratified randomization for clinical trials. *J Clin Epidemiol* 1999;52(1):19-26.
97. Treasure T, MacRae KD. Minimisation: the platinum standard for trials? *BMJ* 1998;317:362-3.
98. Day S. Commentary: Treatment allocation by the method of minimisation. *BMJ* 1999;319:947-8.
99. Chesnut CH, Silverman S, Andriano K, et al. A randomized trial of nasal spray salmon calcitonin in postmenopausal women with established osteoporosis: the Prevent Recurrence Of Osteoporotic Fractures study. *Am J Med* 2000;109(4):267-76.
100. Cummings SR, Chapurlat RD. What PROOF proves about calcitonin and clinical trials. *Am J Med* 2000;109(4):330-1.
101. No author listed. Influence of adherence to treatment and response of cholesterol on mortality in the Coronary Drug Project. *N Engl J Med* 1980;303(18):1038-41.
102. Irvine J, Baker B, Smith J, Jandciu S, et al. Poor adherence to placebo or amiodarone therapy predicts mortality: results from the CAMIAT study. Canadian Amiodarone Myocardial Infarction Arrhythmia Trial. *Psychosom Med* 1999;61(4):566-75.
103. Horwitz RJ, Viscoli CM, Berkman L, Donaldson RM, et al. Treatment adherence and risk of death after a myocardial infarction. *Lancet* 1990;336(8714):542-5.
104. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ* 1999;319(7211):670-4.
105. Bucher HC, Guyatt GH, Cook DJ, et al. Users' guides to the medical literature: XIX. Applying clinical trial results. A. How to use an article measuring the effect of an intervention on surrogate end points. Evidence-Based Medicine Working Group. *JAMA* 1999;282(8):771-8.
106. Fleming TR, DeMets DL. Surrogate end points in clinical trials: Are we being misled? *Ann Intern Med* 1996;125(7):605-13.
107. Sobel B. Biomarkers, surrogates, and evaluation of therapeutic agents: a proposed solution. *Cor Art Dis* 1999;10:339-41.
108. Ericsson CG, Hamsten A, Nilsson J, Grip L, et al. Angiographic assessment of effects of bezafibrate on progression of coronary artery disease in young male postinfarction patients. *Lancet* 1996;347(9005):849-53.
109. Bestehorn HP, Rensing UFE, Roskamm H, et al. The effect of simvastatin on progression of coronary artery disease: the multicenter Coronary Intervention Study (CIS). *Eur Heart J* 1997;18(2):226-34.
110. MAAS Investigators. Effect of simvastatin on coronary atheroma: the Multicentre Anti-Atheroma Study (MAAS). *Lancet* 1994;344:633-8.
111. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet* 1994;344(8934):1383-9.
112. Secondary prevention by raising HDL cholesterol and reducing triglycerides in patients with coronary artery disease: the Bezafibrate Infarction Prevention (BIP) study. *Circulation* 2000;102(1):21-7.
113. Clarke M, Oxman AD (editors). Cochrane Reviewers' Handbook 4.1.1 (updated December 2000). In: The Cochrane Library, Issue 4, 2000. Oxford: Update Software. Updated quarterly. <http://www.cochrane.dk/cochrane/handbook/handbook.htm>
114. Egger M, Smith GD, Altman D (editors). Systematic Reviews in Health Care, 2nd Edn. *BMJ Books*, 2001.
115. Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. *BMJ* 1997;315(7121):1533-7.
116. Smeeth L, Haines A, Ebrahim S. Numbers needed to treat derived from meta-analyses – sometimes informative, usually misleading. *BMJ* 1999;318:1548-51.
117. Tang J-L. Weighting bias in meta-analysis of binary outcomes. *J Clin Epidemiol* 2000;53:1130-36.
118. Egger M, Smith GD. Misleading meta-analysis. (editorial). *BMJ* 1995;310:752-4.
119. Egger M, Smith GD, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629-34.
120. Thompson SG. Systematic Review: Why sources of heterogeneity in meta-analysis should be investigated. *BMJ* 1994;309:1351-5.
121. Naylor CD. Meta-analysis and the meta-epidemiology of clinical research. *BMJ* 1997;315(7109):617-9.
122. Egger M, Schneider M, Smith GD. Meta-analysis: Spurious precision? Meta-analysis of observational studies. *BMJ* 1998;316:140-44.
123. Egger M, Smith GD, O'Rourke K. 1. Rationale, potentials, and promise of systematic reviews. In: Egger M, Smith GD, Altman D (editors). Systematic reviews in Health Care, 2nd Edn. *BMJ Books*; 2001.
124. Mulrow CD. Systematic reviews: rationale for systematic reviews. *BMJ* 1994;309(6954):597-9.
125. Altman DG, Bland JM. Statistics notes: Absence of evidence is not evidence of absence. *BMJ* 1995;311:485.
126. Smith GD, Egger M, Phillips AN. Meta-analysis: beyond the grand mean? *BMJ* 1997;315:1610-14.
127. Sniderman AD. Clinical trials, consensus conferences, and clinical practice. *Lancet* 1999;354(9175):327-30.
128. Greenhalgh T. "Is my practice evidence-based?" (editorial). *BMJ* 1996;313:957-8.
129. Sackett DL, et al. Clinical epidemiology. A basic science for clinical medicine, 2nd edition. Boston: Little Brown; 1991.

(Az internetes címek érvényessége 2001. március 30-án ellenőrizve.)

Szerkesztőségi megjegyzés: A közleményt és a kommentárt a szerzők beleegyezésével a *Ca és Csont* című folyóirat 2001/1. számából vettük át, bízva abban, hogy lapunk olvasói is hasznos ismereteket meríthetnek belőle.

Kommentár

A *Bizonyítékok és a végpontok hierarchiája* című közlemény elolvasását ajánlom mindazoknak, akik képesek arra, hogy adott esetben az orvostudomány eredményeinek tükrében felülvizsgálják gyógyszer-felírási szokásaikat; azoknak, akik betegek kezelésénél inkább a racionális, bizonyítékokra támaszkodó gyógyszerkezelés elveit kívánják követni, mintsem a divat vagy a lukratív promóciós anyagok csábító tanácsait. A dolgozat elkészítése és közzlése abból a szempontból is üdvözölhető, hogy a szerző sok oldalról megvilágítja a bűvös szlogen, a bizonyítékokon alapuló orvoslás (evidence based medicine) jelentését, tartalmát.

A szerző igen nagyszámú irodalmi forrásra hivatkozik, de a közleményből érződik, hogy maga is jól ismeri és nagyra becsüli Cochrane-nek, a téma apostolának híres munkáját.

Szelíd kritikái megjegyzéseim semmit sem vonnak le a közlemény értékéből, inkább csak arra utalnak, hogy a „legvitathatatlanabb” eredmények értékelésénél se felejtkezünk el a vizsgálati rendszer esetleges hibáiról. A korszerű, objektív gyógyszer-vizsgálati eljárások eredményeit ugyanis gyakran nevezünk – fenntartás nélkül – tényeknek. És jaj nekünk, ha ezektől a tényektől nem tudjuk megismerni az igazságot.

A dolgozat bevezető fejezetei egyértelműen azt kívánják alátámasztani, hogy a „késlekedő gyakorlat” (számos klinikai vizsgálat metaanalízisének hiánya) azzal az eredménnyel jár, hogy – különösen krónikus betegségekben – nem vesszük észre a kicsi, ám szignifikáns klinikai változásokat és ezáltal „még több felesleges szenvedést és halál”-t okozunk. Ez az érvelés elfogadható. Úgy gondolom azonban, hogy számos klinikai vizsgálat jól végzett metaanalízise legalább annyira azt a célt is kell hogy szolgálja, hogy végre megállapíthassuk, hogy a kérdéses gyógyszer kevésbé hatásos vagy egyszerűen hatástalan. Ilyen vizsgálati eredményt azonban csak ritkán publikálnak. Személyes tapasztalatom, hogy metaanalízis ritkábban volna végezhető, mint ahogy azt ma megteszik. Még multicentrikus vizsgálatokban is, ahol a vizsgálatok szigorúan azonos protokoll alapján folynak, ugyancsak szigorú monitorozás mellett, nemritkán nehéz a különböző egységekben kapott eredmények együttes értékelése. Ugy látom, hogy gyakran egy kisebb betegszám mellett végzett, jól kontrollált, úgynevezett pivotál vizsgálat eredménye megbízhatóbb, mint számos vizsgálat összemossa.

Elfogadom, hogy a bizonyítékok hierarchiájában a kettős vak, kontrollált, randomizált vizsgálat ered-

ményei a legértékesebbek, de szabad legyen a következőket megjegyezni. Maga a véletlenszerű besorolás is eredményezhet torzításokat, és az eredmények csak akkor fogadhatók el, ha az egyes csoportokba véletlenszerűen besorolt betegek adatai a vizsgálat kezdetén nem különböznek, azaz szigorúan homogén csoportok szerepelnek a vizsgálatban, és ezt statisztikailag ellenőrzik.

A kettős vaktság biztosítása szintén elősegíti, hogy objektív eredményeket nyerjünk. Nagyon ritkán olvasok azonban olyan vizsgálatokról, ahol validálták a kettős vaktság folyamatos fennmaradását a vizsgálat folyamán. (Így például egy β -receptor-blokkoló alkalmazása során akár már 24 óra múlva is könnyű rájönni, melyik szer tartalmazza az aktív anyagot. Elég csak a beteg pulzusszámát figyelni egy könnyebb fizikai megterhelés után.)

Sok esetben etikai szempontból nem lehetséges placebót kapó kontrollcsoport beállítása; a referens készítmények alkalmazásával végzett, úgynevezett non inferiority vizsgálatok értéke, értékelhetősége gyakran megkérdőjelezhető.

Végül megjegyzem, hogy a törzskönyvezés előtt végzett összehasonlító klinikai vizsgálatoknál a beválasztási kritériumok (éppen a homogenitás biztosítása miatt) általában rendkívül szigorúak. Ezért ezek a vizsgálatok soha sem pótolják a törzskönyvezés után a gyakorlati gyógyítómunka során végzett megfigyeléseket.

Elfogadom, hogy amikor egy krónikus betegség kezelésére vagy annak prevenciójára alkalmazott szer esetén apróbb pozitív változásokat akarunk értékelni, a szerzővel egyetértésben a szerényebb eredményeket is komolyan kell venni. Nem szabad azonban arról elfeledkezni, hogy amikor megállapítjuk, hogy például több száz beteg éveken át tartó kezelésével megakadályozzuk néhány betegnél a myocardialis infarctus kialakulását, egyúttal hány betegnél okozunk enyhébb vagy súlyosabb gyógyszer-mellékhatásokat (hepatotoxicitás, vesetoxicitás, csontvelő-károsító hatás, allergiás reakciók stb.).

Összefoglalva: valóban, a bizonyítékokon alapuló orvoslás – ami, úgy érzem, ma már nemcsak egy divatos szlogen, hanem lassan tartalommal töltődik meg – a korszerű orvoslás alapját kell, hogy képezze. A közlemény bőséges bizonyítékát szolgáltatja annak, hogy erre szükség van. Megjegyzéseim inkább arra vonatkoznak, hogy semmit sem szabad fetisizálni, és végső soron egy-egy új gyógyszernek terápiás vagy preventív hatását nagyon sokoldalúan kell értékelni, s ennek egyik, de nem egyedüli eszköze a randomizált, kettős vak, kontrollált vizsgálat, illetőleg e vizsgálatok metaanalízise.

dr. Borvendég János