

Az empirikus orvosi kutatások pár alapvető kérdése

Okozatiság és kutatása az empirikus orvosi vizsgálatokban: kísérlet és megfigyelés

Ferenci Tamás

Óbudai Egyetem, Neumann János Informatikai Kar, Élettani Szabályozások Csoport

Bevezető biostatistikai cikksorozatomban első részében az empirikus orvostudományi kutatások pár alapvető gondolatát tisztázom, igyekezve a különféle kutatási módszereket egyetlen közös gondolkodási keretben elhelyezni. Ezen belül is különös hangsúllyal tárgyalom a kísérlet és a megfigyelés különbségét, kitérve arra, hogy milyen eltérések vannak az így nyerhető információk bizonyító erejében.

(Megjelent az Immunonkológia (ISSN: 2498-6364) folyóirat 2016/1. számában, pp. 20-21.)

A legtöbb orvosi kutatás egyetlen mondatban összefoglalható: arra vagyunk kíváncsiak, hogy valamilyen *expozíció* vajon *hat-e* valamilyen *végpontra*. Első ránézésre ez meglepő állítás lehet, de ha jobban végiggondoljuk, akkor azt látjuk, hogy – a konkrét kutatás egyedi részleteitől megtisztítva, kellően absztrahálva – a séma egy sor esetben ez; csak az „expozíciót” és a „végpontot” kell megfelelően definiálni. Ide tartoznak a szokásos gyógyszerkísérletek (magnöveli a túlélési időt az onkológiai készítmény? – *expozíció*: gyógyszerbevitel, *végpont*: túlélési idő; okoz-e hasmenést, mint mellékhatás, ugyanaz a szer? – *expozíció*: gyógyszerbevitel, *végpont*: hasmenések előfordulása), de ide tartoznak a klasszikusan epidemiológiaiak hívtak kérdések (növeli-e a rákkockázatot a vörös hús fogyasztása? – *expozíció*: vörös hús fogyasztás, *végpont*: rák-előfordulás), sőt, az egészen klinikainak tűnő kérdések is (igaz-e, hogy a császármetszéssel születettek körében több lesz később az 1-es típusú cukorbeteg? – *expozíció*: császármetszéssel születés, *végpont*: T1DM előfordulás).

Tovább gondolkodva ebben a keretben, azt is észrevehetjük, hogy a kérdés megválaszolásának a logikája is legtöbbször ugyanaz: úgy döntünk az *okozatiságról*, hogy megvizsgáljuk, *eltérnek-e az expozíció szerint képezett csoportok a végpontban*. Tovább élnek a kezelt betegek? Többnek lesz köztük hasmenése? Sűrűbben fordul elő vastagbélrák a több vörös húst fogyasztók között? Több 1-es típusú cukorbeteg lesz a császármetszéssel születettek között?

Ezt a döntést több tényező nehezíti. Az egyik, hogy a fenti logika kizárólag abban az esetben helytálló, ha a csoportok tényleg *csak* az expozícióban térnek el. Ekkor – és csak ekkor! – joggal mondhatjuk, hogy a végpontban talált esetleges különbség valóban az expozíciónak (és esetleg a véletlen ingadozásnak) tudható be. Ha azonban nem csak ebben van eltérés, onnantól kezdve nem tudhatjuk, hogy a végpontbeli különbség – ha találunk is ilyet – minek tudható be: az expozíciónak, az egyéb eltérésnek, esetleg a kettő valamilyen keverékének...? Egy nevezetes példa [1] a hormonpótlásos kezelés (HRT): a kezdeti vizsgálatokban összehasonlították az ilyen kezelést kapó és nem kapó nők között a kardiovaszkuláris események előfordulását, és azt tapasztalták, hogy az előbbi csoportban szinte fele olyan gyakoriak a koszorúér-betegségek (CHD). Akkor tehát a HRT véd ez ellen? Sajnos – mint azt így utólag visszatekintve már tudjuk – ez nem igaz, sőt, kissé növeli is a kockázatot. Akkor miért nem ezt láttuk az egyszerű összehasonlításban? A probléma, hogy a HRT-vel kezelt és nem kezelt csoportok nem csak a kezelés tényében tértek el: jellegéből adódóan ezt a kezelést inkább olyanoknak írták fel, akik magasabb szocioökonómiai státuszúak voltak (magasabban iskolázottak, inkább városlakók stb.). Igen ám, csak hogy a szocioökonómiai státusz meg önmagában csökkenti a CHD-rizikót! Hiszen az ilyen nők inkább járnak sportolni, inkább odafigyelnek az egészséges táplálkozásra és így tovább. Végeredményben tehát a kezelésben részesülés ténye nem csak a kezelésben részesülést jelentette, hanem eltérő szocioökonómiai státuszt is, ám míg az egyik növeli, addig a másik *csökkenti* a CHD rizikóját – márpedig egyszerű összehasonlításnál ez a kettő egybemosódott, olyannyira, hogy az utóbbi el is fedte az előbbit. Úgy is megfogalmazhatjuk mindezt, hogy a két változónk (HRT szedés és CHD kockázat) mögött ott volt egy harmadik (szocioökonómiai státusz), ami egyszerre hatott mindkettőre: a HRT-szedésre *pozitívan*, a CHD rizikóra *negatívan*. Ha ezt a harmadik változót is figyelembe vesszük, akkor a *tiszta* hatás a HRT és a CHD rizikó között pozitív, de ha figyelmen kívül hagyjuk, akkor – teljesen fals módon – negatív.

Ez a legalapvetőbb példa arra a jelenségre, amit úgy szoktunk hívni, hogy *confounding* (ritkán használt magyar szóval: zavaró változós hatás); az ilyen harmadik változó neve: *confounder* vagy zavaró változó [2].

Ez az állandó rákfénje minden olyan vizsgálatnak, ahol csak passzíve tudjuk megfigyelni az expozícióban részesülést, de arra nincsen ráhatásunk. Az ilyeneket szokás *megfigyeléses (obszervációs) vizsgálatnak* nevezni. Amikor ilyet olvasunk, *mindig* gondoljunk a confoundingra, és tegyük fel magunknak a kérdést: vajon a csoportok *tényleg* csak a vizsgált expozícióban térnek el...? Biztos, hogy a császármetszéssel születők csak a szülés módjában térnek el a per vias naturales születőktől, és semmi másban? Biztos, hogy a több vörös húst fogyasztók csak és kizárólag a fogyasztott vörös hús mennyiségében térnek el a kevesebb vörös húst fogyasztóktól...? Mert ha netántán nem, akkor onnantól kezdve – még ha találunk is különbséget T1DM, illetve rák előfordulásában – nem lehetünk benne biztosak, hogy azt mi okozza: a vizsgált hatás, az egyéb eltérés(ek), vagy esetleg ez(ek) valamilyen keveréke?

Természetesen, ha eszünkbe jut, hogy mi lehet a potenciális confounder (és le tudjuk azt mérni – pont a szocioökonómiai státusz jó példa arra, ahol ez sem triviális!), akkor mód van arra, hogy ennek hatását *statisztikai eszközökkel* (rétegzés, többváltozós statisztikai modellek alkalmazása, az ún. propensity score matching) *kiszűrjék* [3]. Ha tehát

eszünkbe jut ez a probléma, és nem csak az alanyok HRT-szedését és CHD-eseményeit jegyezzük fel, hanem a szocioökonómiai státuszukat is, akkor mód van arra, hogy – annak ellenére, hogy megfigyelések az adataink! – kiszűrjük a szocioökonómiai státusz hatását statisztikai úton, és így igyekezzük a HRT-szedés és a CHD rizikó közötti valódi kapcsolatra következtetni. Ahogy mondani szokták: megvizsgáljuk a kettő közti kapcsolatot úgy, hogy *kontrollálunk* a szocioökonómiai státuszra.

A probléma, hogy ezt csak akkor tudjuk megtenni, ha tudatában vagyunk annak, hogy valami confounderként viselkedik és azt le is tudjuk mérni. Az egyetlen *univerzális* megoldás erre a problémára, ha képesek vagyunk befolyásolni az expozíciót, ekkor (és csak ekkor!) ugyanis megtehetjük, hogy *véletlenszerűen* tesszük ki neki az alanyokat, márpedig így (és csak így!) biztosítható, hogy az expozíció szerint képezett csoportokban *biztosan* nincsen *semmilyen* szisztematikus – azaz a véletlen ingadozáson túlmenő – különbség az expozíció tényén kívül. Azokat a vizsgálatokat, ahol a kutatók aktívan befolyásolják az expozícióban részesülést szokás *kísérletes (experimentális) vizsgálatnak* nevezni.

Látható tehát, hogy a kísérlet hatalmas előnye az, hogy automatikusan védelmet jelent a confounding ellen: az eszünkbe jutó, véges számú confounder ellen megfigyeléses vizsgálatban is lehet védekezni, de a kísérlet igazi ereje, hogy *automatikusan* védelmet jelent az *összes, végtelen sok* confounder ellen: azok ellen is, amiket nem tudunk mérni, sőt, azok ellen is, amik eszünkbe sem jutnak! (Természetesen a kísérlet végzése ehhez csak szükséges, de nem elégséges feltétel: lehet rosszul végezni, ami miatt ez az előny mégsem realizálódik, ezzel később részletesebben fogok foglalkozni.)

Akkor miért nem végzünk mindig kísérletet? A probléma – azon túl, hogy sokszor drága, idő- és szervezésigényes –, hogy sok esetben nem kivitelezhető: aligha lehet randomizáltan császármetszésnek alávetni anyákat, vagy véletlenszerűen vörös húst „fogyasztatni” emberekkel 30 évig, és utána megnézni, hogy hány rákos lesz közöttük... Ilyen esetekben csak a megfigyelés segít – annak minden nehézségével együtt. (Talán többen ismerik azt a cikket [4], melyben a szerzők, nyilván a bizonyítékokon alapuló orvoslás túlkapásain való gúnyolódásként is, azt vizsgálták, hogy az ejtőernyő jelent-e segítséget a „gravitációs challenge” leküzdésében, ha az ember nagy magasságban kiesik egy repülőgépből. Véresen komoly írás, pontosan definiálva, hogy milyen trauma súlyossági score a végpont, milyen adatbázisokat kerestek, milyen szűkítésekkel és így tovább... majd arra konkludáltak, hogy egyetlen randomizált, kontrollált kísérletet sem végeztek, mely ezt vizsgálta volna, úgyhogy igazából nem lehet tudni, hogy segít-e, ha van rajtunk ejtőernyő, amennyiben kiesünk egy repülőgépből!)

Fontos hangsúlyozni, hogy a fentiekben *bizonyítóerőbeli különbségekről* van szó – de ettől még ezek a vizsgálatok „egy ligában játszanak”! Olyannyira, hogy sok jól tervezett megfigyelés helyettesíthet egy kísérletet is. Mindig *bizonyítékok hierarchiájában* érdemes gondolkodni: nem arról van szó, hogy van kutatás ami „rossz”, más meg „jó”, csak *eltérő a bizonyítóerejük* [5]. Nem tudok róla, hogy bárki kísérletben – pláne placeboval kontrollált kísérletben – megvizsgálta volna, hogy appendicitisben hatásos beavatkozás-e az appendectomia! (Pedig nem lenne nehéz, csak meg kellene várni, amíg a kezelt csoportból szinte mindenki gyógyultan hazamegy, a kontrollcsoportból pedig szinte mindenki meghal szeptikus sokkban perforált vakbéllel.) Fontos, hogy nem egyszerűen azért nem hajtanak végre ilyen kísérletet, mert „nyilvánvaló”, hogy mi lenne az eredménye, hanem

épp azért, mert számos megfigyeléses bizonyíték áll rendelkezésre, melyek összességében kiváltják az igényt a kísérlet végzésére. Mint az előbbi példa is mutatja, a kísérletek esetében – mivel ott *beavatkozunk* a betegek életébe – ez etikai kérdés is: kísérlet végzésének csak akkor van helye, ha *valódi bizonytalanság* van az expozíció hatását illetően [6].

Végezetül fontos megjegyezni két dolgot. Az egyik, hogy a bizonyítóerőnek ez nem két fokozata: a megfigyelésen és a kísérleten belül is vannak eltérő (adott esetben nagyon eltérő!) erejű vizsgálatok. Az orvosi kutatások kritikus értékelésével mérhetjük ezt fel egy adott cikket olvasva; erre hangsúlyosan ki fogok térni cikksorozatom későbbi írásaiban is. A másik, hogy valójában van egy harmadik szint is: az előbbi két szintbe tartozó vizsgálatok – lehetőleg szisztematikus – *összegyűjtése* (szisztematikus review) és az *eredményeik aggregálása* (meta-analízis) [7]. Ez egyrészt lehetővé teszi többféle kérdés vizsgálatát, például a közismert problémáét, hogy a gyógyszercégek el tudják titkolni a nekik nem tetsző eredményt hozó kutatásokat, másrészt növeli a kapott eredmény megbízhatóságát. Kis költői túlzással mondhatjuk, hogy ma már mindenre és az ellenkezőjére is lehet egy perc alatt kutatást találni; ezért különösen fontos a „*bizonyítékok összessége*” szemlélet, melyre én is ki fogok a későbbiekben térni.

Hivatkozások

- [1] D. A. Lawlor, G. Davey Smith, and S. Ebrahim. “Commentary: The hormone replacement–coronary heart disease conundrum: is this the death of observational epidemiology?” In: *International Journal of Epidemiology* 33.3 (2004), pp. 464–467. DOI: 10.1093/ije/dyh124. URL: <http://ije.oxfordjournals.org/content/33/3/464.short>.
- [2] S. Greenland and H. Morgenstern. “Confounding in Health Research”. In: *Annual Review of Public Health* 22.1 (2001), pp. 189–212. DOI: 10.1146/annurev.publhealth.22.1.189. URL: <http://www.annualreviews.org/doi/abs/10.1146/annurev.publhealth.22.1.189>.
- [3] R. McNamee. “Regression modelling and other methods to control confounding”. In: *Occupational and Environmental Medicine* 62.7 (2005), pp. 500–506. DOI: 10.1136/oem.2002.001115. URL: <http://oem.bmj.com/content/62/7/500.short>.
- [4] G. C. Smith and J. P. Pell. “Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials”. In: *British Medical Journal* 327.7429 (2003), pp. 1459–1461.
- [5] D. Evans. “Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions”. In: *Journal of Clinical Nursing* 12.1 (2003), pp. 77–84. ISSN: 1365-2702. DOI: 10.1046/j.1365-2702.2003.00662.x. URL: <http://onlinelibrary.wiley.com/doi/10.1046/j.1365-2702.2003.00662.x/full>.
- [6] B. Freedman. “Equipoise and the Ethics of Clinical Research”. In: *New England Journal of Medicine* 317.3 (1987), pp. 141–145. DOI: 10.1056/NEJM198707163170304. URL: <http://www.nejm.org/doi/pdf/10.1056/NEJM198707163170304>.

- [7] M. Borenstein, L. Hedges, J. Higgins, and H. Rothstein. *Introduction to Meta-Analysis*. Wiley, 2011. ISBN: 978-1-119-96437-7.