



"PÁR PERC TUDOMÁNY"

Hogyan érdemes orvosbiológiai cikkeket olvasni?

Sematikus iránytű a klinikusokat elárasztó tudományos cikkek tengerén való eredményes navigációhoz.

Dr. Tél Bálint, PhD

Semmelweis Egyetem, Gyermekgyógyászati Klinika, Bókay utcai részleg

A PubMed adatbázisa alapján a biomedicinai közlemények éves száma 2010 óta megduplázódott, a 2000-es évekhez képest közel háromszorosára nőtt, ami a 2025-ös évben naponta átlagosan több, mint 5000 új orvostudományi cikk megjelenését jelentette. E többnyire megbízható információtengernek köszönhetők az elmúlt évtizedek áttörései, az új gyógymódok és biztonságosabb kezelési protokollok. Ugyanakkor az óriási mennyiség nem jelent egyforma tudományos minőséget. Ez az összefoglaló egyszerűsített iránymutatást ad ahhoz, hogy hogyan mérjük fel kritikusan egy olvasott cikk alkalmazhatóságát és korlátait.

MILYEN CÉLLAL OLVASOM A CIKKET?

Egy tudományos cikk súlya nem feltétlenül állandó, elsődlegesen attól függ, hogy amikor olvassuk, milyen klinikai/preklinikai kérdésre keressük éppen a választ, arra tud-e válaszolni. Ha érdeklődésből olvasunk, elég lehet a szerzők által megfogalmazott célkitűzésekből kiindulni. Ilyenkor is fontos azonban, hogy tisztán kiderüljön, mi volt a vizsgált klinikai kérdés, ellenkező esetben felmerülhetnek kérdések a szerzők módszertani rigora kapcsán. **A jól körülírt klinikai kérdés** megfogalmazásához legtöbbször a **PICO** sémát használják, melyhez az alábbi négy szempontot kell pontosan definiálnunk:

- P** (patients/population/problem) – a vizsgálni kívánt célpopulációt (magába foglalja az összes vizsgálati kárt);
- I / E** (intervention / exposure) – a vizsgálni kívánt hatást (mely lehet beavatkozás vagy expozíció);
- C** (comparison) – az összehasonlítási alapot (pl. placebo, másik kezelés vagy expozíció hiánya);
- O** (outcome) – és a hatást mérő, előre rögzített kimenetelt.

[TOVÁBB OLVASOM](#)

A MÓDSZERTANI RIGOR (VALIDITÁS) VIZSGÁLATA I. – MEGFELELŐ-E A VIZSGÁLATI MÓDSZER?

A különböző klinikai kérdések vizsgálatához különböző módszerek illenek: egy gyógyszer hatékonyságáról a legmegbízhatóbb evidenciákkal például egy randomizált kontrollált vizsgálat szolgálhat, de ugyanez a vizsgálat – a szigorúan beválogatott, kísérleti kezelésben is részesülő populáció miatt – nem alkalmas az adott betegség természetes lefolyásának vagy egy tünet gyakoriságának megbecslésére, ehhez kohorsz- vagy keresztmetszeti vizsgálatra lenne szükség. Ahhoz, hogy felmérjük az olvasott publikáció módszertani validitását, először ellenőrizzük, hogy a központi klinikai kérdés megválaszolásához mennyire volt alkalmas az adott vizsgálati elrendezés!

- Terápia hasznossága? → randomizált kontrollált vizsgálat
- Etiológia? Káros hatás? → kohorsz vagy eset-kontroll
- Prognózis? → inception kohorsz
- Diagnosztikus pontosság? → keresztmetszeti vagy prospektív kohorsz
- Prevalencia / incidencia? → keresztmetszeti vagy prospektív kohorsz

	Randomizált kontrollált vizsgálat	Kohorsz vizsgálat	Eset-kontroll vizsgálat
P	A vizsgálatba bevont résztvevők.	A vizsgálat indulásakor még betegségmentes populáció.	A populációt a már kialakult kimenetel határozza meg.
I/E	A vizsgálni kívánt aktív kezelés / beavatkozás.	Az "exponált" státusz.	Az eseteknél retrospektíven vizsgált tényező (expozíció?).
C	Placebo vagy szokásos ellátás.	Nem exponált csoport.	Korábbi expozíció hiánya vagy eltérő mértéke.
O	Időbeli követés betegség/kimenetel	Időbeli követés betegség/kimenetel	Retrospektív betegség/kimenetel

[TOVÁBB OLVASOM](#)

[TOVÁBB OLVASOM](#)

A MÓDSZERTANI RIGOR (VALIDITÁS) VIZSGÁLATA II.

Az egyes vizsgálati módszerek megfelelő módszertani alkalmazásának megvannak a saját szabályai, és minden vizsgálati módszer más- és más hibalehetőségeket, torzításokat tartalmazhat. Éppen ezért az elmúlt két évtizedben minden egyes vizsgálati típushoz jelentek meg szigorú, a publikációhoz irányt mutató útmutatók (**"reporting guideline"**), melyeket az **"Equator Network"** fog össze. Ezek ismerete nagy segítség a cikkek validitásának felméréséhez, a részletes leírásuk azonban meghaladja ezen iránytű kereteit. Az alábbiakra azonban minden esetben érdemes odafigyelni:

- A szerzők előre meghatározott (optimálisan publikált) vizsgálati protokollt követtek?
- Csatolták-e a cikkhez a megfelelő reporting guidelinet (Equator Network) ellenőrző listáját?
- Tisztán kiderülnek-e a beválogatási, kizárási kritériumok, az esetdefiníciók, alcsoportok, küszöbértékek, stb.?

[TOVÁBB OLVASOM](#)

[TOVÁBB OLVASOM](#)

A MÓDSZERTANI RIGOR (VALIDITÁS) VIZSGÁLATA III. – MI BEFOLYÁSOLHATTA AZ EREDMÉNYEKET?

Minden vizsgálati módszer eredendően okozhat torzításokat (**"bias"**), mert sohasem végtelen esetszámú, teljesen a való életnek megfelelő mintákat vizsgálunk. A torzítások nem elkerülhetők, de odafigyeléssel minimalizálhatók. Olvasói szempontból kiemelten fontos felismerése, és annak megítélése, hogy a torzítás (és mértéke) befolyásolhatta-e az eredményeket. A torzítások egy része függ a vizsgálati módszertől (pl. a nem megfelelő randomizációból eredő torzítás), vannak azonban univerzálisan minden esetben felmerülő torzítások:

- Kiválasztási torzítás:** A nem egyenletes beválogatás (pl. egy centrum egy rendelőjének esetei) okozza. Ha a vizsgálati csoportok nem mindenben hasonlóak, a kapott eredmények fakadhatnak ezen különbségekből is.
- Információs torzítás:** A nem egyformán gyűjtött adatokból származik (pl. előben a rendelőben versus utólag az ambuláns lapról felvett adatok).
- Hiányzó adatok:** Ha nem minden résztvevőtől rendelkezünk adatokkal egy érték vagy kimenetel kapcsán, vagy ez nem egyformán oszlik el a csoportokban. Mivel statisztikai módszerekkel részben kezelhető a torzítás, annak megléte és a módszerének pontos leírása alapvető elvárás.
- Zavaró tényezők (confounding):** Olyan harmadik tényezők, amelyek nem a vizsgálat hatásból eredően befolyásolják a kimenetelt, de összefüggésben állnak a vizsgált expozícióval, a vizsgálati felállással (pl. dohányzás, szociális körülmények). Különböző módon „kontrollálhatók” (pl. csak nem dohányzók beválogatása, beválogatás (matching), vagy statisztikai korrekció (többváltozós modell)), melynek explicit leírása elvárt.
- Publikációs torzítás és érdekütközés:** A vizsgálatok során kapott teljesen negatív/nem szignifikáns eredmények ritkábban kerülnek közlésre. A közölt eredmények teljességét tovább befolyásolhatja a finanszírozó (pl. cégek) és más összeférhetetlenségek. Emiatt az egyedülálló, sorból kilógó „pozitív” cikk önmagában óvatosan értelmezendő.

[TOVÁBB OLVASOM](#)



STATISZTIKA RÖVIDEN

– MIT JELENT AZ EREDMÉNY?

A klinikai vizsgálatok többsége hipotézisvizsgálaton alapul, amelynek legismertebb jelzőszáma a P-értékkel alátámasztott statisztikai szignifikancia. Ez azonban pusztán egy matematikai módszer eredménye, és önmagában nem azonos a klinikai jelentőséggel. Fontos tudni, hogy a P-értéket erősen befolyásolja az elemszám, nagy mintában egy klinikailag jelentéktelen különbség is matematikailag szignifikánssá válhat, egy kis mintában viszont a valós hatás akár el is veszhet. Emiatt lett alapvető elvárás az előzetes elemszámbebecslés, ahol a vizsgálatot végzők előre kiszámolják, hány résztvevő kellene egy klinikailag releváns hatás megbízható kimutatásához. Az alultervezett vizsgálat nemcsak elszalaszthatja a valódi hatásokat, hanem paradox módon a szignifikánsnak talált eredményeknek rendszerint túlbecsüli a nagyságát. A szignifikancia tehát sosem egyenlő a klinikai relevanciával. Gyakorlati szempontból többet mondanak a „klinikai mutatók”: a hatás nagysága, az effektusméret („effect size”) és a 95%-os konfidencia-intervallum, a relatív kockázat, vagy a szükséges kezelésszám (NNT), és diagnosztikus vizsgálatoknál a pozitív/negatív prediktív értékek (PPV/NPV). Az eredmények átnézésekor érdemes figyelni a torzításokat korrigáló módszerek meglétére, a többszörös tesztelés megfelelő korrekciójára, és óvatosan kezelni az utólagos alcsoport-elemzések eredményeit.

[TOVÁBB OLVASOM](#)

CIKKEK A MESTERSÉGES INTELLIGENCIA KORÁBAN

A mesterséges intelligencia (AI) és a nagy nyelvi modellek (LLM) használata robbanásszerűen terjed az orvosi irodalomban is, részben vizsgálati **módszerként**, részben pedig **„társszerzőként”**. Az AI-módszereket alkalmazó publikációk kritikus olvasása néhány speciális szempontot is igényel, ezeket részletesen tárgyalják az erre kialakított *reporting guideline*-ok, mint a TRIPOD+AI, a TRIPOD-LLM és a STARD-AI. Ezekben szereplő két legfontosabb kérdés:

- Történt-e prospektív, **külső validáció** (más populáción, más centrumban), vagy csak a betanító adathalmazon mért, önmagát igazoló teljesítményt látunk?
- Kizárható-e az **adat-szivárgás** („leakage”), amikor a tesztalmaz információja már a modell tanításába is beszivárgott, mesterségesen növelve annak pontosságát?

Az AI mint „írásasszisztens” használata utólag nehezen állapítható meg, bár bizonyos szóhasználat és tipográfiai jegyek alapján sejthető. Emiatt a legújabb cikkekben már elvárt, hogy a szerzők tételesen nyilatkozzanak arról, hol és mihez vettek igénybe AI-segítséget.

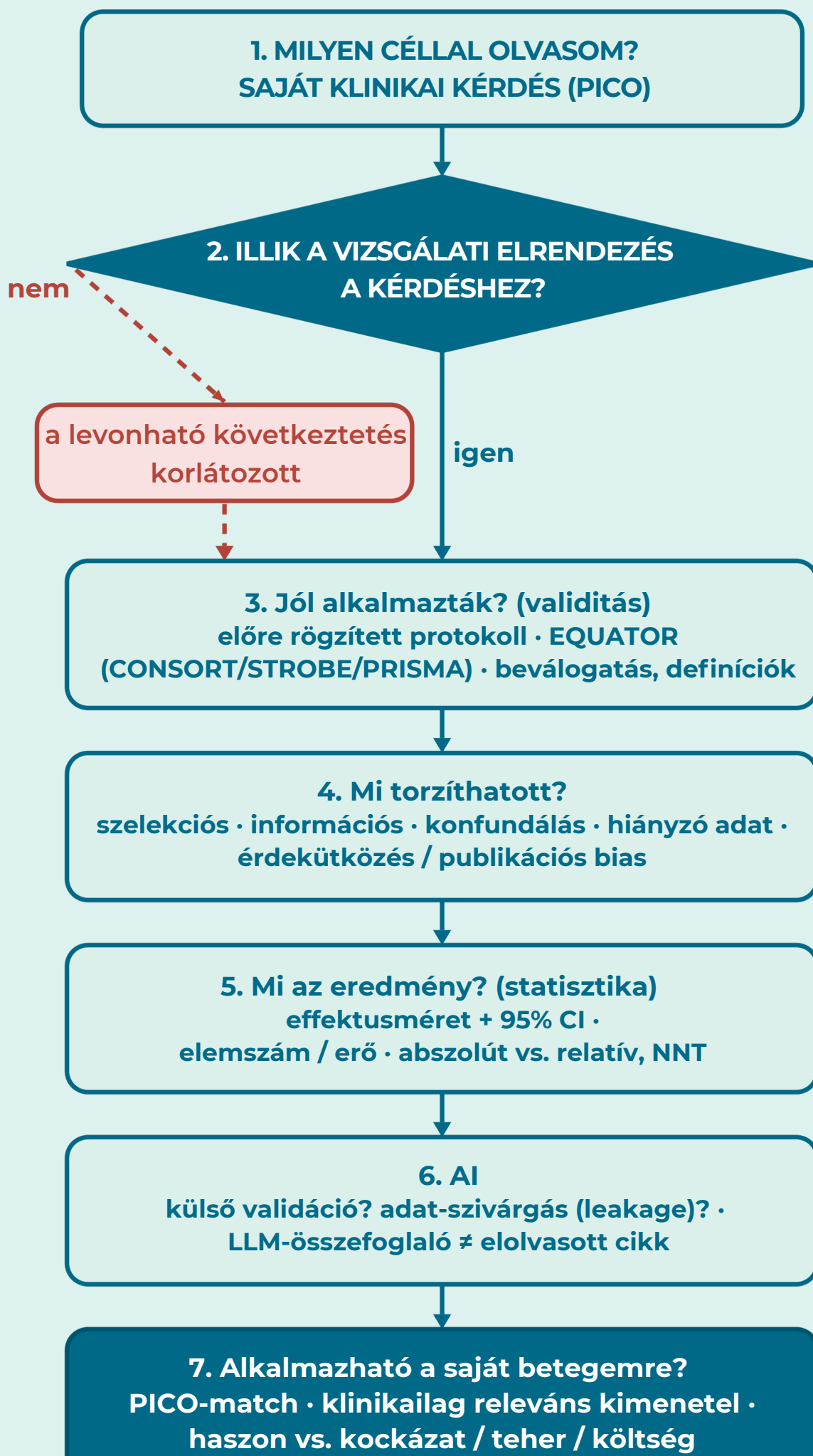
Végül fontos megemlíteni, hogy bármennyire vonzó a cikk helyett egy „LLM-bot” (pl. ChatGPT) által készített összefoglaló olvasása, az nem helyettesíti a cikk alapos elolvasását. A modellek – főleg az egyszerűbbek – figyelmen kívül hagyhatják az eddigiekben részletezett kritikai lépéseket vagy kihagyhatnak lényeges szempontokat (pl. a torzításokat), sőt nem létező adatokat vagy hivatkozásokat is megidézhetnek („hallucináció”).

[TOVÁBB OLVASOM](#)

ALKALMAZHATÓ A SAJÁT BETEGEMRE/KLINIKAI KÉRDÉSEMRE?

Még a módszertanilag kifogástalan, statisztikailag meggyőző eredmény sem feltétlenül alkalmazható a saját betegünkre. A külső validitás, alkalmazhatóság felméréséhez érdemes visszatérni a PICO-hoz, és összevetni a vizsgálat populációját a saját klinikai helyzetünkkel, mennyire van átfedésben az életkor, a társbetegségek, az ellátási környezet? A szigorú beválogatás miatt a vizsgálati minta gyakran „egészségesebb” a hétköznapi betegnél, így a valós hatás kisebb, vagy akár teljesen eltérő is lehet. Fontos továbbá, ha alkalmazható is a mért kimenetel, a gyakorlatban valóban releváns-e a betegeink számára, és az eredmény alkalmazása egyáltalán elérhető és kivitelezhető-e? A végső döntés mérlegeléséhez pedig azt kell felmérni, hogy a várható haszon arányban áll-e a kockázatokkal, a megterheléssel és a költséggel, továbbá illeszkedik-e a beteg értékrendjeihez?

ÖSSZEFOGLALÓ ÁBRA



AI-NYILATKOZAT:

A cikk írása során a szerző az alábbi elemekhez vett igénybe részleges AI asszisztenciát (Anthropic Claude, Fable 5 és Opus 4.8):

- Célzott irodalomkutatás
- Ábra-generálás
- Szöveg tömörítés, nyelvi ellenőrzés