

KLINIKAI VIZSGÁLATOK ÉRTÉKELÉSÉNEK SZEMPONTJAI

How to evaluate clinical trials?

Torday László¹, Végh Éva²

¹SZTE ÁOK Onkoterápiás Klinika, Szeged

²Egyesített Szent László és Szent István Kórház és Rendelőintézet, Budapest

ÖSSZEFOGLALÓ – Egy klinikai vizsgálat adatait mindig nagy odafigyeléssel kell szemlélni, és mindenképpen javasolt az adatok részletekbe menő áttanulmányozása. Először minden esetben érdemes meghatározni azt, hogy a vizsgálati adatok milyen evidenciszinten adják meg az eredményt, és ezek alapján milyen ajánlás tehető. Meg kell vizsgálni a vizsgálat típusát, elrendezését, betegszámát, statisztikai erejét, randomizációját, stratifikációját, a karok kiegyensúlyozottságát, a végpontok és a követési idő megfelelését az adott klinikai kérdés megválaszolására. A kimutatott statisztikai szignifikancia mellett az adatok klinikai relevanciáját, a kimutatott eltérések valós nagyságát, és az időfüggő változókra gyakorolt jelentőségét mindenképpen érdemes át gondolni, mert ez szabja meg az adott vizsgálat hatását a valós klinikai gyakorlatra. A „kreatív vizsgálati tervek” beköszöntött korszakában kifejezett odafigyelést és kritikus hozzáállást igényel az, hogy meggyőződhessünk arról, hogy a vizsgálat során nem követtek-e el, nem követhető-e el olyan tipikus torzítás, amely a vizsgálat valós eredményeit megváltoztathatná.

Kulcsszavak: bizonyítékon alapuló orvoslás, evidenciaszintek, klinikai vizsgálat, végpontok, biostatistika, hibák

SUMMARY – Data from a clinical study should be always viewed with great care, and is recommended to examine and evaluate the data in-depth. First of all should be determined the evidence level of the trial's results and what recommendations can be made based on them. The study's type and the study design, the number of participating patients, statistical power, randomisation, stratification, the balance between study arms, the endpoints and the adequacy of follow-up time should be examined. In addition the importance of the observed statistical significance, the clinical relevance of the data and the real magnitude of the detected differences, and the time-dependent variables are worth to consider because these factors can determine the impact of the given clinical trial on the every-day clinical practice. In the era of the “creative study plans” a critical attitude is necessary to be convinced that the study does not contain any typical bias that could influence the real value and actual results of the trial.

Key words: evidence based medicine, levels of evidence, clinical trial, endpoint, biostatistics, bias

Az elmúlt időszakban számos olyan, onkológiai betegek bevonásával végzett klinikai vizsgálat eredményét publikálták, melyek megítélése a tudományos közösség berkein belül vagy a hatóságok részéről nem volt egyértelmű. Ezekben az esetekben leírva vagy kongresszusi megbeszélések során számos esetben vitatták a vizsgálati terv megalapozottságát, az eredmények interpretálását, illetve a szignifikáns különbségek je-

lentőségét. A klinikai vizsgálatok lefolytatása ennek ellenére rendkívül fontos lépés az alapkutatás és a klinikai gyakorlat között. A specifikus kérdésekre választ adó klinikai vizsgálatok biztosítják a bizonyítékokon alapuló orvostudomány azon evidenciáit, melyekre a klinikai gyakorlat tökéletesedése alapul.

A vizsgálatok értékelése az evidenciaszinteknek megfelelően

Egy leközölt klinikai vizsgálat elsődleges értékét az szabja meg, hogy milyen szintű evidenciát szolgáltat a vizsgálat által felvetett kérdés megválaszolására, és

Levelező szerző:

dr. Tímár József, Semmelweis Egyetem, Klinikai Központ, 2.sz. Patológiai Intézet, 1091 Budapest, Üllői út 93.
Telefon: +36-1-2156921, email: jtymar@gmail.com

ebből milyen fokozatú ajánlás tehető a tesztelt terápiás opcióra a klinikai használat vonatkozásában. Erre a bizonyítékokon alapuló orvoslás több mint két évtizede kialakított (1) szabályrendszerén belül pontos meghatározások léteznek, de az onkológiai gyakorlatban az ASCO által meghatározott evidenciaszint-meghatározást és ajánlási fokozatrendszert használjuk. Az ASCO rendszerét az ESMO is átvette, és irányelveiben konzisztensen alkalmazza. Általánosságban kijelenthető, hogy egy vizsgálat annál értékesebb, minél magasabb szintű evidenciát képes szolgáltatni. Emiatt, ha egy vizsgálat értékét meg kívánjuk határozni, elsősorban azt kell tisztázni, hogy az adott kérdést illetően a vizsgálat milyen típusú vizsgálat, és ehhez kapcsolódóan milyen evidenciaszinttel rendelkezik.

A vizsgálatok típusai

Kísérletes vizsgálatok

A kísérletes vizsgálatokban a vizsgálók egy beavatkozást alkalmaznak, és annak hatását elemzik. A kísérletes vizsgálatok általában randomizáltak. Fontos megjegyezni, hogy nem minden kontrollált vizsgálat randomizált, de minden randomizált vizsgálat kontrollált.

Randomizáció – A randomizáció során a betegeket véletlenszerűen osztják el a vizsgálati karok (vizsgálati szer/protokoll vagy kontroll) között. Ez az eljárás csökkenti a vizsgálói és a mintából eredő hibát.

Stratifikáció – A betegeket jellemzőik alapján válogatják (például a betegség kiterjedése, prognosztikai faktorok), abból a célból, hogy a vizsgálati karok kiegyensúlyozottak legyenek, és így az eltérő prognosztikai karakterisztikából fakadó kiegyensúlyozatlanság nem befolyásolja a vizsgálat kimenetelét. A stratifikáció (rétegzés) alkalmazása lehetővé teszi azt is, hogy a

vizsgálók megfelelő alcsoport-analíziseket is végre tudjanak hajtani.

Randomizált kontrollált vizsgálatok (randomized controlled trial, RCT)

A beválasztott betegeket random módon két vagy több kezelési karba sorolják. Az egyik karon a vizsgálati szert/protokollt alkalmazzák, a másik karon pedig az addigi arany standard kezelést.

Kontrollált vizsgálatok (controlled clinical trial, CCT)

Hasonló az RCT-hez, de a betegek allokálása a vizsgálati és kontroll csoportba nem random történik. Ez megemeli annak a hibának a lehetőségét, hogy a hasonló jellemzőjű vizsgálati személyek egy csoportba kerülve befolyásolhatják a végső eredményeket.

Megfigyeléses/beavatkozás nélküli vizsgálatok (observational/non interventional studies)

Elsősorban kockázati faktorok, diagnosztikai tesztek, vagy kezelések eredményét figyelik meg anélkül, hogy befolyásolnák a történeteket (2). Ezek a vizsgálatok fontos adatokat biztosíthatnak az elmúlt években globálisan felértékelődő gyógyszer-biztonságosságához. A prospektív, beavatkozással nem járó vizsgálatok alatt észlelt nemkívánatos események jelentésére világszerte és hazánkban is (3) szigorú farmakovigilanciái szabályok vonatkoznak. Bizonyos esetekben ezek a vizsgálatok retrospektívek, mert az adatok a már megtörtént eseményeken alapulnak. A legtöbb saját beteganyagfeldolgozás is ebbe a kategóriába esik.

Kohorszvizsgálatok (cohort studies)

A kohorsz olyan vizsgálati csoport, mely tagjai egy adott tulajdonságban megegyeznek, és egy adott időtartamon át követik őket. A kohorszvizsgálatok során az egyik kohorszot kiteszik valamilyen behatásnak, melyet aztán összevetnek egy hasonló kohorsszal, akik nem voltak kitéve a vizsgált változónak.

Eset-kontroll vizsgálatok (case-control studies)

Itt a vizsgálók már meglévő betegadatokat használnak arra, hogy kialakítsanak egy betegcsoportot, akiknél jelen vannak a vizsgált jellemzők, illetve egy olyan csoportot, akiknél pedig nem, majd ezek adatait összevetik.

Evidenciaszintek és az evidenciaszinteken alapuló ajánlások fokozatai

A bizonyítékon alapuló orvoslás (EBM: evidence based medicine, magyar rövidítéssel BAO) negyedszázada került be az orvoslásba (1) és alakította ki a bizonyítékon alapuló egészségügyi ellátások rendszerét. A klinikai vizsgálatok eredményeire támaszkodva, azokat szintetizálva alakítják ki a különböző szakmai szervezetek az irányelveiket.

Az evidenciaszinteket különböző rendszerekben határozzák meg, ilyen például az oxfordi Centre for Evi-

Legfontosabb megállapítások

- Egy publikált klinikai vizsgálat elsődleges értékét az adja meg, hogy milyen szintű evidenciát szolgáltat a vizsgálat által felvetett kérdés megválaszolására.
- Az onkológiai vizsgálatokban a gyakran használt elsődleges és másodlagos végpontok definíciójának ismerete nélkülözhetetlen
- Még korrektül tervezett, végrehajtott és feldolgozott klinikai vizsgálat esetén is a közölt statisztikai szignifikancia (p-érték) mellett a klinikai relevanciát is értékelni kell.
- A vizsgálatokról szóló közlemények értelmezésénél a klinikai vizsgálati elrendezés, az elsődleges, másodlagos és exploratív végpontok megismerése mellett fontos a statisztikai próbák és eredményeik megértése, az esetleges hibák, torzítások észlelése is.

1. táblázat. A vizsgálati típusokhoz tartozó evidenciaszintek és az evidenciaszintekből levezethető ajánlások fokozatai.

Az evidenciaszintek	
I	Az evidencia metaanalízisből vagy több, jól tervezett kontrollált vizsgálatból származik, melyek esetében kevés az álpozitív és álnegatív hiba (nagy statisztikai erővel rendelkeznek).
II	Az evidencia legalább egy jól tervezett kísérletes vizsgálatból származik vagy olyan randomizált vizsgálatokból, ahol sok az álpozitív és álnegatív hiba (alacsony erővel rendelkeznek).
III	Az evidencia jól tervezett, kvázi kísérletes vizsgálatokból, például nem randomizált, kontrollált, egycsoportos, pre- és poszt-, kohorsz-, idő- vagy párosított eset-kontrollos sorozatokból származik.
IV	Az evidencia jól tervezett, nem kísérletes vizsgálatokból, mint komparatív és korrelációs deskriptív vagy esettanulmányokból származik.
V	Az evidencia esetismertetésekéből és klinikai esetekből származik.
Az evidenciaszinteken alapuló ajánlás fokozatai	
A	I. típusú evidencia létezik, vagy a II., III. vagy IV. típusú vizsgálatok eredményei konzisztensek.
B	II., III. vagy IV. szintű evidencia létezik, és az eredmények általában konzisztensek.
C	II., III. vagy IV. szintű evidencia létezik, de az eredmények inkonzisztensek.
D	Nincs vagy csak kevés szisztematikus tapasztalati evidencia van.

dence Based Medicine (CEBM) rendszere [4], amely a terápia/prevenció/etiológia/károk mellett a diagnózisra és a prognózisra is kialakította saját rendszerét. Ettől különböző rendszert használ a daganatos betegségek kezelésére vonatkozó irányelveiben az NCCN (National Comprehensive cancer Network) [5].

A vizsgálati típusokhoz tartozó evidenciaszinteket és az evidenciaszintekből levezethető ajánlások fokozatait az **1. táblázat** tartalmazza.

Az evidencia bázisa (3. rendű)

Az evidencia mennyisége – Elsősorban az ajánlás alapjául szolgáló döntést alátámasztó klinikai vizsgálatok száma, de a vizsgálatokon részt vevő betegek száma a megfigyelt esemény gyakoriságához képest ugyanilyen fontos lehet (más szavakkal, milyen volt a vizsgálat ereje). Az evidencia mennyisége az addigi klinikai vizsgálatokat inkorporáló metaanalízissel is meghatározható.

Az evidencia szintje – Ezt általában a kérdés megválaszolására rendelkezésre álló legjobb szintű klinikai vizsgálat adja meg. A klinikai vizsgálatok típusainak a vizsgálati felépítésből fakadó inherens torzításai miatt adott evidenciaszinthez tartozó vizsgálatok szisztematikus áttekintése, vagy metaanalízise sem emeli meg az evidenciaszintet még akkor sem, ha az áttekintés, vagy metaanalízis színvonala esetleg kivételes (ez alól csak a jól tervezett II. evidenciaszintű vizsgálatok összefoglalói, vagy metaanalízisei a kivételek).

Az evidencia minősége – Azt mutatja, hogyan folytatták le a vizsgálatot a torzítások elkerülésének érdekében, azaz, miképpen történt a vizsgálati alanyok kiválasztása, allokálása a vizsgálati karokba, menedzselése, és követése, és miképp határozták meg a vizsgálati végpontokat.

Az adatok konzisztenciája

A konzisztencia meghatározza, hogy az evidencia alapjául szolgáló adatok a rendelkezésre álló vizsgálatokon belül mennyire konzisztensek. Ez a mutató azt tükrözi, mennyire valószínű az, hogy a vizsgálati eredmények eltérő beteganyagon, eltérő vizsgálati elrendezések és körülmények esetén reprodukálhatóak. Ezt például egy metaanalízis esetén a heterogenitásra vonatkozó statisztikai analízis tükrözi, melyet azonban éppen a felsorolt tényezők befolyásolhatnak a leginkább. A nem randomizált vizsgálatok a nagyobb torzítási valószínűség miatt nagyobb hatást mutathatnak, de ennek ellenére ezek a vizsgálatok fontosak lehetnek a randomizált vizsgálatok eredményeinek megerősítésében vagy megkérdőjelezésében.

A klinikai hatás mértéke

Ez a betegek részére elérhető potenciális előnyök mérőszáma, mely függ attól, hogy a nyert evidencia mennyire releváns egy klinikai kérdés megválaszolásában, mennyire fontos a rendelkezésre álló kezelési opciók mellett a betegek kezelése szempontjából, milyen statisztikai erővel mekkora a kimért hatás, mennyi ideig kell kezelni a betegeket a hatás eléréséhez, és megvan-e a megfelelő kiegyenlítettség a kockázatok és a haszon között.

Generalizálhatóság

Ez a mérőszám azt adja meg, hogy a klinikai vizsgálatból nyert evidencia mennyire alkalmazható egy valóban létező betegcsoportnak a valós klinikai gyakorlatban kezelt tagjaira. Mivel a klinikai vizsgálatokban kezelt betegcsoport jellemzői specifikusak, így a figyelembe vett komponensek közé kell, hogy tartozzon a vizsgálatban kezelt betegek életkora, származása, a

beválasztáskori prognosztikai tényezők, s a rendelkezésre álló egészségügyi infrastruktúra és kezelési lehetőségek is.

A vizsgálatok fázisai az evidenciaszinteknek megfelelően

A klinikai vizsgálatok fázisai és az azonos fázison belüli típusok

A gyógyszerek, illetve kombinációk klinikai vizsgálatokon belül zajló tesztelése meghatározott fázisokon keresztül zajlik, általában kisebb fázis I-es vizsgálattal indul, s nagy, randomizált, fázis III. vizsgálatban teljesedik ki. A fejlesztési fázisok célja eltérő, melyek megértése alapvető egy vizsgálat értékelése szempontjából (6).

Fázis I. vizsgálat – Sokszor az első humán vizsgálati lépés, miután a laboratóriumi vizsgálatok és állatkísérletek alapján valószínűsíthető, hogy a gyógyszer hasznos lehet a daganat kezelésében. A vizsgálat célja a biztonságos dózistartomány megtalálása (MTD: maximálisan tolerálható dózis), a mellékhatások felmérése (DLT: dózist limitáló /korlátozó/ toxicitás), esetleg korai hatásossági információk nyerése. A résztvevők száma alacsony, általában 20–80 beteg. Az onkológiai korai fázisú klinikai vizsgálatok speciálisak, hiszen míg más indikációjú gyógyszerek alkalmazásánál az I. fázisú vizsgálatokban egészséges önkéntesek vesznek részt, az onkológiai vizsgálatokban – a szerek előre várható toxicitása miatt – erre gyakorlatilag nincs lehetőség. Ezért szigorú bevásztási és kizárási kritériumok alapján vonják be az onkológiai betegségük mellett kevés egyéb kísérőbetegségben szenvedő betegeket a korai onkológiai vizsgálatokba, s gyakran összevont fázis I-II vizsgálatokra kerül sor (7).

Fázis II. vizsgálat – Ha a gyógyszer sikeresen átmegy az első fázison, következik ez a szakasz. A fázis II. vizsgálat a gyógyszer/kezelés farmakológiai hatása alapján kiválasztott indikációban végzett vizsgálata. Céljai közé tartozik a vizsgálati készítmény terápiás hatásának igazolása, a dózis-hatás összefüggés megállapítása, az optimálisnak tartható egyszeri és napi dózis meghatározása, a mellékhatások megfigyelése. Ennek során meghatározzák, hogy a gyógyszer milyen daganatok esetében alkalmazható, milyenek a mellékhatásai, ezeket miképpen lehet menedzselni, mi lehet a legjobb dózis, és ami a legfontosabb, az eredmények alapján van-e annyira hatásos a gyógyszer, hogy fázis III. vizsgálatokban teszteljék tovább. Ebben a fázisban a bevont betegszám is nagyobb: 100–300 beteg. A fázis II. vizsgálatok lehetnek II/A, illetve II/B vizsgálatok. A II/A vizsgálatok célja a gyógyszer dozírozásának pontos karakterizálása, míg a II/B vizsgálatok a hatásosságot tesztelik az előírt dózisban. Vizsgálati elrendezés szempontjából egy fázis II. vizsgálat lehet egykarú (lényegében ez egy „case series”, általában III. szintű evidenciát ad), illetve randomizált kontrollált (általában II. szintű evidenciát ad). Az utóbbi

eseten az új gyógyszert/kezelést egy már meglévővel (ami az esetek túlnyomó részében az addigi kezelési arany standard) hasonlítják össze. Placebokontrollos vizsgálatot csak olyan esetben etikus végezni, ha nem áll rendelkezésre megbízható és elfogadott kezelés az adott betegcsoport részére. Amennyiben az új gyógyszer/kezelés jobb vagy legalább annyira hatásos, mint a régi, akkor kerülhet sor a fázis III. vizsgálatokra. A randomizált, kontrollált fázis II. vizsgálatokkal elérhető evidenciaszint általában II, mert ezek a vizsgálatok a kis betegszám miatt általában csekély statisztikai erővel rendelkeznek. Ezt kell szem előtt tartani, amikor egy fázis II. vizsgálat bizonyító erejét vagy a klinikai gyakorlatra tett hatását vizsgáljuk.

Fázis III. vizsgálatok – Ebben a vizsgálati fázisban az új kezelést az addigi legjobb elérhető kezeléssel (más néven: standard kezelés: a legelterjedtebben használt kezelési modalitás egy adott daganattípus kezelésében) vetik össze. Ennek során pontosan meghatározzák az új kezelés hatásosságát, és igazolják a vizsgálati készítmény biztonságos alkalmazhatóságát, kontrollált, randomizált vizsgálati elrendezésében. A gyógyszerek törzskönyvezéséhez leggyakrabban ilyen fázisú vizsgálat szükséges. A fázis III. vizsgálatok általában nagy statisztikai erejű, jól tervezett, jól vezetett vizsgálatok. Sokszor multicentrikusak, mely azt jelenti, hogy a vizsgálat különböző vizsgálóhelyeken és/vagy több országban is folyik. A bevont betegek száma igen nagy lehet, akár 300–3000 is. A fázis III. vizsgálatokkal nyert adatok evidenciaszintje attól függ, hogy a vizsgálat megerősítő vizsgálat-e, vagy egy terápiás indikációban az első. A második esetben az elérhető evidenciaszint II., míg az első esetben a vizsgálat által nyújtott evidencia a már meglévő adatok megerősítése révén I. szintű is lehet.

Fázis IV vizsgálatok – Több ezer, a rutin orvosi gyakorlatban kezelt betegen végzett adatgyűjtés, a már törzskönyvezett, forgalomba hozatali engedéllyel rendelkező gyógyszert az elfogadott indikációban, az elfogadott dózistartományban, az elfogadott alkalmazási előírásnak megfelelően felhasználva. Céljuk a gyógyszer valós klinikai körülmények közti hasznának és kockázatának pontosabb felmérése, a kockázat-haszon arány meghatározása, az adagolás pontosítása, s a ritkábban előforduló mellékhatások felismerése. A fázis IV. vizsgálatok által szolgáltatott evidencia a vizsgálati terv függvénye, mert ezek a vizsgálatok lehetnek randomizált klinikai vizsgálatok, de lehetnek akár beavatkozással nem járó, megfigyeléses vizsgálatok is. Rendkívül fontos megjegyezni, hogy annak ellenére, hogy a megfigyeléses fázis IV. vizsgálatokkal nyert adatok az inherens torzítás miatt nem biztos, hogy pontosak, de kiválóan alkalmasak arra, hogy a meglévő randomizált, kontrollált klinikai vizsgálatokból nyert adatokat megerősítsék vagy azokat megkérdőjelezzék.

Próbavizsgálatok (Pilot studies) – A próbavizsgálatot azelőtt folytatják le, mielőtt a nagy fázis III. klinikai vizs-

gálatot elkezdének. A pilot study a fővizsgálat kisebb verziója, melynek fő célja annak tesztelése, hogyan működnek a tervezett vizsgálat fő alkotóelemei. A pilot study már segíthet abban, hogy a fő vizsgálati kérdésre választ kapjunk, és a vizsgálói csapat a pilot study adatait felhasználhatja a fővizsgálat adatainak analízise során. A próbavizsgálatokból származó adatok a kis betegszám és csekély statisztikai erő miatt nem haladják meg a II. evidenciaszintet.

Bioekvivalencia vizsgálatok – Általában egy vizsgálóhelyen lebonyolított klinikai vizsgálatok, amelyek a generikus gyógyszerfejlesztés fontos részét képezik. Céljuk a referencia- (originális) és tesztkészítmény koncentráció-idő görbéjének meghatározása, a farmakokinetikai paraméterek vizsgálata [8].

Bioszimiláris gyógyszereket fejlesztő vizsgálatok – Az onkológiában széles körben alkalmaznak biológiai rendszerekben előállított makromolekulákat. Az originális készítmény védettségének lejárta után fejlesztik a követő (biohasonló, bioszimiláris) termékeket, amelyeknél (a generikus készítményekkel ellentétben) nem teljesülhet a teljes kémiai azonosság követelménye [8, 9]. A bioszimiláris gyógyszerek fejlesztésére az amerikai és az európai gyógyszerhatóságok (EMA, FDA) szigorú irányelveket dolgoztak ki [10–13].

A klinikai vizsgálatok analízisében használt biostatistikai fogalmak

Mielőtt egy klinikai vizsgálat adatait közlik, a statisztikusok analizálják az eredményeket. Összehasonlítják a kapott eredményeket a vizsgálok által vártakkal, és meghatározzák az adatokra vonatkozó statisztikai szignifikanciát azaz azt, hogy az adatok mennyiben jelentenek valós különbséget, és mennyire igaz az, hogy nem a véletlen okozta az különbségeket.

Egy klinikai vizsgálat eredményeinek közlése során számos általánosan használt fogalom adatait írják le.

Erő (power) – Az a mérték, mely elsősorban a vizsgálat nagyságán alapszik, hogy egy vizsgálat milyen valószínűséggel fog statisztikailag szignifikáns eredményt produkálni. Minél több beteget vonnak be a vizsgálatba, annál nagyobb lesz az erő, mely a vizsgálati és a kontroll kar közötti eltérést detektálni képes.

A p-érték (p value) – Az a mérték, mely meghatározza annak az esélyét, hogy egy klinikai vizsgálat eredménye nem a véletlen folytán keletkezik, hanem valós különbséget jelez a vizsgálati és a kontroll kar között. Tipikus esetben a 0,05-öt, mint a p-érték felső határát határozzák meg, mint a klinikai vizsgálatok eredményének statisztikai szignifikanciájának meghatározóját.

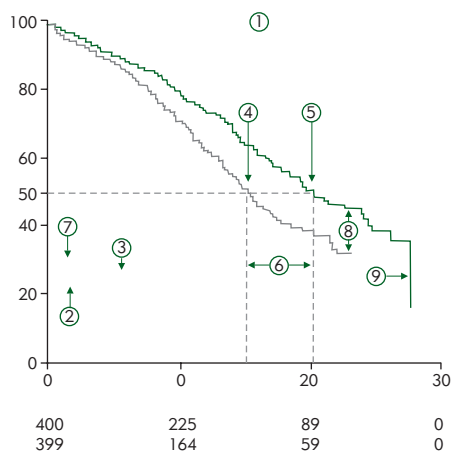
Konfidenciaintervallum, megbízhatósági tartomány (confidence interval, CI) – Az értékek olyan tartománya, azaz intervalluma, amelybe a vizsgált paraméter valós értéke előre láthatóan beleesik. Az intervallum korrelál a valószínűség szintjével, amely általában

95% [95% CI]. Ez azt jelenti, hogy a vizsgáló 95%-ban biztos abban, hogy a vizsgált paraméter valós értéke ebben a tartományban helyezkedik el. A szűk, illetve tág konfidenciaintervallum (megbízhatósági határ) értékes információt ad arról, hogy mennyire pontosak az eredmények.

Relatív kockázat (Hazard ratio, HR) – A detektált esemény (például progresszió) bekövetkeztének relatív kockázata a vizsgálat időtartama és a követési periódus alatt az egyik vizsgálati karon a másikkhoz képest. Ha a HR 1, a kockázat mindkét karon megegyezik. Ha a HR <1, azt mutatja, hogy az egyik vizsgálati karon a kockázat kevesebb (tipikusan a vizsgálati karon). Ha a HR >1, azt jelenti, hogy az egyik karon a kockázat nagyobb. A HR az $(1-HR)/HR$ formula használatával segít meghatározni a vizsgált végpont teljes javulását, például ha a progressziómentes túlélések (PFS-ek) HR-ja 0,53, a PFS teljes javulása $(1-0,53)/0,53$, azaz 89%. A HR szintén használható annak meghatározására, hogy a vizsgált végpont bekövetkeztének kockázata mennyivel csökken. Például, ha a PFS-ek HR-ja 0,53, akkor az azt jelenti, hogy a progresszió kockázatának csökkenése 1-HR, azaz 47%. A HR mellé általában megadják a 95% CI-t is mely azt jelzi, hogy a HR igazi értéke 95%-os biztonsággal hová esik.

A Kaplan–Meier-görbe értelmezése

A Kaplan–Meier (KM) túlélési görbe grafikus ábrázolása egy adott időtartam alatt történő olyan túlélés valószínűségének, melynek definiált kezdőpontja általában a diagnózis időpontja vagy a kezelés kezdete, illetve a randomizálás időpontja, míg végpontja egy meghatározott esemény bekövetkezése, mint például a beteg halála vagy a betegség progressziója. A Kaplan–Meier-becslő az egyik legjobb módja annak, hogy meghatározzuk a kezelés megkezdésétől eltelt meghatározott időpillanatban életben lévő betegek arányát [14]. Ez az adott időpillanathoz tartozó Kaplan–Meier-becslők (estimatorok) berajzolásával történik. A KM-görbe megjelenése lépcsőzetes, függőleges eséssel az időintervallum végén. Különböző okokból a megfigyelési periódus gyakran azelőtt véget ér, mielőtt minden beteg esetében bekövetkezne a meghatározott esemény, ezeket a betegeket cenzorált betegeknek nevezzük. Ha a beteg cenzorálttá válik, nem tudjuk, hogy a várt esemény bekövetkezik-e, s ha igen, mikor, de az adataikat az analízishez felhasználják. A beteg cenzorálását egy piciny grafikai jelölés (általában egy pici függőleges vonal) jelzi a vízszintes vonal felett abban az időpillanatban, amikor a cenzorálás történt. A KM-görbe három feltételezés köre épül: a cenzorált betegek túlélési kilátásai ugyanolyanok, mint akik követése tovább folyik, a korábban illetve később beválasztott betegek túlélési esélyei megegyeznek, és az esemény az adott időn belül bekövetkezik. A második feltételezés különösen fontos a hosszú beválasztási



1. ábra. A Kaplan-Meier görbe analízise

1. A medián követési idő (median follow-up) az az időtartam, ameddig a vizsgálati populáció 50%-át követték.
2. A HR <1 azt jelenti, hogy a vizsgálati karon a kockázat kisebb.
3. A HR 95% CI: alsó határérték 0,64; felső határérték=0,88.
4. A kontrollkar medián PFS-e 15,6 hónap
5. A vizsgálati kar medián PFS-e 20,3 hónap
6. Különbség a két kar medián PFS-e között
7. A p értéke nagy statisztikai szignifikanciát mutat
8. A PFS-különbség egy adott időpillanatban
9. A görbe meredek esése az analízis céljára elérhető betegek kis számából fakad

*Néhány klinikai vizsgálatban a medián PFS-t nem érik el az adatok analízise pillanatában. Ebben az esetben a hatásossági adatokat gyakran mint HR adják meg. Ha Kaplan-Meier-görbével ábrázolják a klinikai vizsgálat adatait, a HR adja a leginkább pontos adatot a kezeléssel elért teljes különbségről, mert a HR a teljes görbe kumulatív adatait reprezentálja.

periódusú klinikai vizsgálatok esetében, mert a kezelési lehetőségek megváltozása megváltoztathatja a túlélési esélyeket is. Két KM-görbe közötti túlélési időkülönbségek kimutatására gyakran használják a lograng statisztikai hipotézis próbát (log-rank test), de ez a teszt nem képes több változót kezelni abból a célból, hogy elválassza a kezelés hatását más, a túlélést befolyásoló tényezőktől. Az utóbbi cél elérésére a Cox-féle arányos kockázat modellt (Cox proportion hazards model) használják. A Cox-modellben a kockázat mértékét minden változóra kiszámolják, ez a relatív kockázat (hazard ratio, HR). Ha a HR >1, azt jelenti, hogy a vizsgált változó a vizsgálati alany számára emeli a kockázatot, míg ha a HR <1, ez csökkent kockázatot jelent (1. ábra).

A relatív kockázat (HR) értékelése

A Cox-modellből származó relatív kockázat nem ad információkat a megfigyelt események bekövetkeztét jellemző időpontokra. Ha a relatív kockázat azt mutatja, hogy az onkológiai kezelés eredményes, ez általában azt jelenti, hogy a túlélési végpontig tartó idő hosszát a kezelés növeli. Azonban az is megfigyelhető, hogy a

HR mértéke kevesebb vagy több lehet, mint a kezelés valódi hatása a medián végponti időtartamokra. Ez különösen olyan görbék esetén igaz, ahol a görbék meredeken esnek egymással párhuzamosan. Emiatt egy vizsgálat esetén a HR mellett mindig meg kell határozni azt, hogy mi a medián végponti időtartamok aránya (medián arány, median ratio: MR). Ez a kezelt kar meghatározott medián végponti idejét (pl. teljes túlélés, OS vagy PFS) a kontroll kar meghatározott medián végponti idejével elosztva határozható meg. Amennyiben ez az érték kicsi, úgy a kezelésnek nincs klinikailag és számszakilag jelentős hatása. Amennyiben ez az érték például kettő, az azt mutatja, hogy a kezelés megduplázza a vizsgált végponthoz tartozó időtartamot. Tehát nem a HR hordozza az információt arra vonatkozóan, hogy hogyan változnak a vizsgált végpontok medián értékei, emiatt egy kezelés terápiás előnyének meghatározása esetén nagy hibákat lehet véteni, ha csak a HR-t vesszük számításba, mert a kedvező hatás nem csak a HR-en múlik, hanem nagyban függ a túlélési görbék alakjától is. A HR a túlélési görbék vertikális axisán mért előnyt mutatja (y-tengely), míg a medián időközön alapuló jellemzők (például MR) a horizontális axison reprezentált előnyt (x-tengely). Ez abból fakad, hogy a HR azt mutatja, hogy egy kezelési csoportba allokált vizsgálati személy esélye mekkora a másikkhoz képest, hogy elérje a meghatározott végpontot. Tehát a HR annak ellenére, hogy pontosan megadja a halálozás vagy progresszió esélyének csökkenését és a vizsgálati adatok megbízhatóságát, a túlélési statisztika adatai azok, amelyek a klinikailag és a beteg számára legrelevánsabb adatokat hordozzák.

A p-érték jelentősége

Nullhipotézis – A p értékének meghatározásához először fel kell tételnie, hogy nincs valódi különbség a két vizsgált kezelés között (14). Ez a feltételezés a nullhipotézis. Ezt követően ki kell számolni, hogy abban az esetben, ha ez a feltételezés, azaz a nullhipotézis igaz, mekkora az esélye annak, hogy az a különbség, amit majd megfigyelünk, csak a véletlen műve abban az esetben, ha a feltételezés igaz. Ez a mérőszám a p értéke. A p értéke tehát azt mutatja, hogy ha valóban nincs különbség a két kar között, akkor mekkora az esélye annak, hogy a klinikai vizsgálatban látott hatás az akkora, mint amilyenek megfigyelik. Ha a p értéke kicsi, akkor nem valószínű, hogy az adatokat a véletlen produkálta, s elutasítható az a nullhipotézis, hogy a két kezelés közt nincs differencia, azaz, elfogadható, hogy a két kezelési kar között van különbség. Ha a p értéke nagy, a megfigyelt különbség oka nyilvánvalóan a véletlen és nem utasítható el az az elgondolás, hogy nincs különbség a két kar között. Általában a 0,05 alatti p-értékek gondolhatók kicsinek. Ebben az esetben 1:20-hoz az esélye annak, hogy a megfigyelt

különbséget a véletlen okozza és nincs valódi különbség, és mondható, hogy a klinikai vizsgálat eredményei szignifikánsak. Kisebb p -értékeket, például a $p < 0,01$ -ot sokszor magasan szignifikánsnak tartjuk, ami azt jelenti, hogy a megfigyelt különbség esélye 1:100, ha feltételezzük, hogy a két kezelés között nincs differencia.

Az 1. és 2. típusú hibák – Ha a konfidenciaintervallumok és a p értékének vizsgálata után kiderül, hogy a kezelési karok közt a különbség statisztikailag szignifikáns, ebből azt a különbséget vonják le, hogy a két kezelés között különbség van. Ellenben csak azért, mert nem valószínű, hogy ekkora különbséget a véletlen produkál, az nem jelenti azt, hogy azt nem produkálhatja véletlen. Tehát például az 1:20 szignifikanciaszinten ($p < 0,05$) mért különbség is lehet a véletlen eredménye. Ezt a hibát 1-es típusú hibának nevezzük.

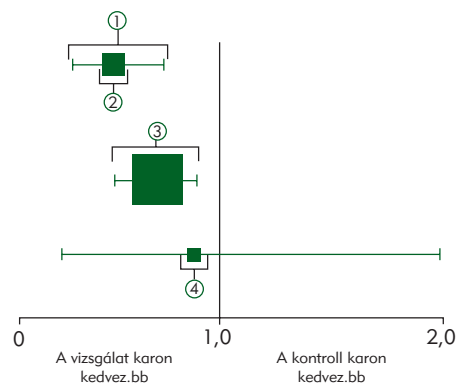
A másik fajta hiba az, amikor egy nem szignifikáns eredményből arra következtetnek, hogy nincs a kezelésnek hatása, amikor valójában pedig van. Ez a 2-es típusú hiba. Az az általános félreértés, hogy a nem szignifikáns jelentése megegyezik a hatástalansággal. A nem szignifikáns konfidencia intervallum csak azt jelenti, hogy a megfigyelt különbség alapján úgy tűnik, nincs valódi differencia a két kar között. Azonban az utóbbi lehetőség nem zárható ki, ami azt jelenti, hogy ha nem találnak szignifikáns kezelési hatást, az nem azt jelenti, hogy nem található szignifikáns kezelési hatás.

A statisztikai szignifikancia és a klinikai szignifikancia különbsége

A statisztikai szignifikancia és a jelentős eredmény közö gyakran kerül egyenlőségjel. Ez az interpretáció egyik nagyon fontos kelepceje. A szignifikancia tesztelése csak azt mutatja meg, hogy a klinikai vizsgálat által produkált adat különbözik-e az új és a referenciakontroll kezeléseket esetén. Ha a két kezelés között különbség mutatkozik, az nem feltétlenül jelenti azt, hogy ez a különbség valóban fontos vagy jelentős. Egy nagy klinikai vizsgálat statisztikailag szignifikánsnak találhat meglehetősen kicsi eltéréseket is, ahol külön kell elbírálni, hogy ezek az eltérések valós klinikai szignifikanciát jelentenek-e. Minden esetben el kell végezni a szignifikáns különbségek ilyen típusú analízisét is, amely arra koncentrál, hogy a hatás valós nagysága (nem pedig a szignifikancia nagysága) az, amely végeredményben számít.

Az alcsoport-analízisek jelentősége

Miután egy klinikai vizsgálat befejeződött, a teljes populáción belül alcsoport-analízisek végezhetőek további részletesebb információ nyerése céljából. Meghatározható, hogy a vizsgált kezelés miként hat egy meghatározott közös jellemzővel leírható betegeket



2. ábra. A „forest plot” adatai

1. A vízszintes vonalak kezdete és vége a konfidenciaintervallum alsó és felső határértékeit jelzi
2. A négyzet helye mutatja a HR értékét, a négyzet nagysága pedig a betegpopuláció relatív nagyságát
3. Ha a HR az 1,0-et reprezentáló vonal bal oldalára esik, az azt mutatja, hogy a vizsgálati kar eredményei a kontroll karhoz képest kedvezőbbek. A nagy kocka azt mutatja, hogy ebben az esetben a betegek viszonylag nagy csoportját analizálták. A szűk konfidenciaintervallum (CI) azt mutatja, hogy az eredmény nagyon megbízható. Az eredmény statisztikailag szignifikáns, hisz a CI nem metszi az 1,0 tengelyt.
4. Bár a HR alapján úgy tűnik, hogy a vizsgálati kar előnyösebb, de a kocka kis mérete azt mutatja, hogy az analizált alcsoport mérete kicsiny, és a nagy CI jelzi, hogy az eredmény viszonylag nem megbízható. Az eredmény statisztikailag nem szignifikáns, hisz a CI metszi az 1,0 tengelyt.

esetében, például idősek vagy kiterjedt betegségben szenvedők. Vizsgálható az is, hogy a teljes populáción látott klinikai előny megtartott-e a különböző beteg típusok esetében, vagy esetleg ez az előny csak egy specifikus alcsoport kezelésével elért sikernek köszönhető-e. Az alcsoport-analízisek lehetnek az adatok feldolgozása előtt előre meghatározottak (prospektív analízis), de végrehajthatóak az első adatfeldolgozás után is (retrospektív analízis). Az előre eltervezett analízis eredményei sokkal megbízhatóbbak, de mindkét fajta alcsoport-analízis eredményeit alaposan meg kell vizsgálni a klinikai vizsgálat adataival együtt, a helytelen következtetések levonásának veszélye miatt. Fontos megjegyezni, hogy a legtöbb klinikai vizsgálatot nem úgy tervezik, hogy minden alcsoport esetében mutassa a szignifikáns eltérést. Emiatt az alcsoport-analízisek eredményeit sokszor csak előzetes adatoknak tartják, függetlenül attól, hogy az prospektív vagy retrospektív volt.

Az alcsoport-analízisek eredménye

Az alcsoport-analízisek eredményeit grafikusán „forest plot” formájában adják meg, amely megmutatja a vizsgált alcsoportok relatív nagyságát, az esetükben tapasztalt klinikai előny mértékét a konfidenciaintervallummal együtt (**2. ábra**).

A végpontok jelentősége

Az elsődleges és másodlagos végpontok kiválasztása

Egy klinikai vizsgálat során a hatásosságát és a biztonsági profilt bizonyos előre meghatározott végpontokkal vagy kimeneteli paraméterekkel mérjük, a vizsgálati tervnek ezek meghatározására kell irányulnia. A végpontok lehetnek klinikai végpontok, mint például a túlélés, vagy helyettesítő végpontok (surrogate endpoints), melyekről azt gondoljuk, hogy előre jelezhetik a megfelelő klinikai végpontokat.

Az elsődleges végpont az a kulcsfontosságú mutató, melynek kapcsán a klinikai hatásosságát értékelik. A vizsgálok által kiválasztott elsődleges végponttól függ a vizsgálat sikeres lefolytatásához szükséges betegszám, így az elsődleges végpontot már a vizsgálat megkezdése előtt meg kell határozni. A másodlagos végpontok másfajta mutatók, melyek további, potenciálisan értékes információt adhatnak. A klinikai vizsgálati protokollnak a másodlagos végpontokat is specifikálni kell, hogy növekedjék az esélye annak, hogy ezen végpontok statisztikai analízise megalapozott legyen.

Az onkológiai vizsgálatokban az elsődleges végpontok kiválasztásánál számos faktort kell figyelembe venni (15). Például, mi az a kezelés által nyújtott előny klinikailag legjelentősebb mutatója, mely a betegség adott stádiumában és adott betegpopuláción a leginkább befolyásolja a klinikai döntéshozatalt? Elfogadható időn belül lefolytatható-e a vizsgálat? Vannak olyan végpontok, melyek hosszabb követési periódust igényelnek, ezek megnyújthatják a vizsgálat idejét ahhoz, hogy megfelelő eredményt lehessen nyerni. Be lehet-e választani a vizsgálat lefolytatásához a megfelelő számú beteget? Ha olyan végpontot választanak, mely a statisztikailag szignifikáns eredmény eléréséhez nagyobb vizsgálatot követel, akkor gondot jelenthet a megfelelő ütemű és mennyiségű betegbeválasztás **(2. táblázat)**.

Teljes túlélés (OS) – A korai 80-as években az FDA elhatározta, hogy a válaszráta nem lehet a gyógyszerbefogadás kizárólagos alapja, mert az nem ad megfelelő információt a gyógyszer valós túlélési és klinikai hasznáról, és egy gyógyszer befogadásához a túlélés vagy a betegek tüneteinek a javulását kell bizonyítani. Az FDA manapság is a teljes túlélést fogadja el, mint a legmegbízhatóbb és preferált végpontot az onkológiai vizsgálatokban. A teljes túlélés megkérdőjelezhetetlen és nem befolyásolja vizsgáloi hiba, így a kockázatok és az előnyök könnyen értékelhetőek. A legfontosabb hátránya, hogy hosszú követést igényel. További problémát jelent, ha a kontroll kar beteget a vizsgálati karra kerülnek betegségük progresszióját követően (cross over), ekkor a túlélési előny eltűnhet. A második vonalas kezelések szintén megemelik a zavaros eredmények esélyét, emiatt, ha a teljes túlélés helyett helyettesítő végpontot használnak, értékelni kell az első vonalas kezelés hatását a második és harmadik vonalas kezelések elindítása előtt.

Teljes válaszráta (ORR) – Az ORR a daganatellenes aktivitás tiszta evidenciája. Az ORR a parciális és komplett válaszok összege. Mivel az adat 2-3 hónapon belül realizálható, az ORR helyettesítő végpontként elfogadott olyan randomizált kontrollált vizsgálatokban, melyek felgyorsított befogadást szolgálnak. 2002-ig az onkológiai gyógyszerek 47%-át ez alapján fogadták be. Az ORR hátránya, hogy nem képes direkt mérni a klinikai előnyt, és nem jelzi előre a kezelés teljes kedvező hatását. Számos problémát jelentenek a RECIST szerinti értékelések gyengeségei is, ezek elkerülése céljából a tumor denzitásának, térfogatának mérése, valamint kvantitatív funkcionális molekuláris képalkotó eljárások segíthetnek.

Betegségmentes túlélés (DFS) – Akkor nagyon hasznos, ha a várható hosszú túlélés miatt nem célszerű az OS mérése. Elsősorban adjuváns klinikai vizsgálatokban használatos. 1990 óta a DFS helyettesítő végpontként elfogadott az adjuváns emlő- és coloncarcinoma vizsgálatokban.

Progressziómentes túlélés – A PFS hasonló egyesített végpont, mint a DFS, ebből fakad előnye is a TTP felett, mert a halálozások detektálásában elkerüli azt a hibát, ami a TTP-analízisekben a meghalt betegek cenzorálásából ered. A PFS hátránya, hogy gyakori radiológiai és más tumorfelmérést igényel, így sokszor a progresszió pontos ideje nem ismert, sőt elsősorban nyílt vizsgálatok esetén sokszor előfordul értékelési hiba is.

Progresszióig eltelt idő (TTP) – Abban különbözik a PFS-től, hogy a progresszió előtt bekövetkező halálozásokot cenzorálják. A TTP-t nem befolyásolja a kezelési kar második vonalas kezelése során bekövetkező esetleges váltása (cross over). Számos klinikai vizsgálatban a TTP definíciója nem pontos, emiatt alkalmazása megfelelő óvatosságot igényel. Hátránya, hogy minél hosszabb az idő két tumorfelmérés között, a TTP annál hosszabb tud lenni.

A betegek által szolgáltatott eredmények (patient reported outcomes, PROs) – Ez általában az életminőség (QRQ). Ezeket az adatokat nagyon nehéz használni, mert alapjuk szubjektív adatközlés, és emiatt a hibák minimalizálása céljából pontos karakterizálásukra kettős vak, placebokontrollált vizsgálatok kívánatosak.

A helyettesítő végpontok értelmezése

A kemény klinikai végpontok, mint például a teljes túlélés (OS, az onkológiai klinikai vizsgálati végpontok arany standardja) vagy daganatspecifikus túlélési markerek használata lelassíthatja az erőfeszítéseket, hogy az új fejlesztések a piacra, így a betegekhez jussanak. Például egy olyan klinikai vizsgálat, mely a prosztatarák kezelésére szolgáló gyógyszert teszti, és a teljes túlélést használja mint daganatspecifikus túlélési végpontot, körülbelül 10 évig tart, míg a vizsgálat indulásától az első beszámoló megtörténik. Amióta

2. táblázat. Általánosan használt végpontok

Végpont	Definíció	Előny	Hátrány
Teljes túlélés (Overall survival, OS)	A randomizáció időpontjától* a bármely okból bekövetkezett halálozásig eltelt idő	A kezelés adta direkt klinikai előny általánosan elfogadott mérőszáma. Könnyen és pontosan mérhető	Nagyobb betegszámot és hosszabb követést igényel a kezelési csoportok közötti statisztikailag szignifikáns különbségek detektálására. A vizsgálati karok közötti keresztezés (cross-over) és a vizsgált kezelést követő terápiák jelentősen befolyásolhatják. A nem daganatspecifikus halálozás is beletartozik
Progressziómentes túlélés (progression-free survival, PFS)	A randomizáció időpontjától* a betegség progressziójáig vagy a beteg haláláig eltelt idő	Kis betegszámot igényel és a teljes túlélésnél rövidebb követési időt. A stabil betegségeket (SD) is magába foglalja. Nem befolyásolja az ágak közötti keresztezés (cross over) vagy a vizsgált kezelést követő terápiák. Általában objektív és kvantitatív értékelésen alapul	Helyettesítő végpontként (surrogate marker) történő megerősítése és használata néhány esetben nehéz lehet. Nem mindig mérhető precíz módon. Definíciója változhat a klinikai vizsgálatok között. Gyakori radiológiai/egyéb értékelést kíván. Az értékelési időpont szempontjából a karok kiegyenlítettége fontos
Progresszióig eltelt idő (time to progression, TTP)	A randomizáció időpontjától* az objektív tumorprogresszióig eltelt idő (a halálozás nincs benne)		
A kezelés sikertelenségéig eltelt idő (time to treatment failure, TTF)	A randomizáció időpontjától* az objektív tumorprogresszióig eltelt idő (a halálozás nincs benne)	Azokban az esetekben hasznos, amikor a kezelés toxicitása valószínűleg ugyanannyira jelentős, mint a betegség progressziója (pl. például allogén őssejttranszplantáció)	Nem tudja a hatásszámot megfelelően differenciálni a toxicitástól
Eseménymentes túlélés (event-free survival, EFS)	A randomizáció időpontjától* a betegség progressziójáig, halálozásig, vagy bármilyen ok miatt bekövetkező terápia megszakításig eltelt idő (pl. toxicitás, betegpreferencia, vagy új kezelés indítása progresszió detektálása nélkül)	Hasonló a PFS-hez; nagyon toxikus kezelések értékelésében fontos	Egy új kezelés indítása szubjektív. A hatóságok nem ajánlják, mert kombinálja a hatásszámot, a toxicitást és a betegek kilépését a vizsgálatból
A következő kezelésig eltelt idő (time to next treatment, TTNT)	A primer kezelés végétől a következő kezelés végéig eltelt idő.	Gyógyíthatatlan betegség esetén fontos végpont a betegek számára	Használata nem elterjedt mint elsődleges végpont. Függsz a lokális kezelési gyakorlattól
Teljes válaszaránya (overall response rate, ORR)	Azon betegek aránya, akik esetében a tumortömeg az előzetesen meghatározott mértékben csökkent.	Egykarú vizsgálatok esetén is értékelhető. Kisebbszámot igényel, és a túlélést elemző vizsgálatokkal szemben hamarabb értékelhető. A mért hatás csakis a gyógyszernek köszönhető, nem pedig a betegség természetes lefolyásának.	A gyógyszer hatásszámának nem átfogó mutatója
A válasz időtartama (duration of response, DoR)	A dokumentált tumorválasztól a betegség progressziójáig eltelt idő.		

*Nem randomizált vizsgálatok esetén a klinikai vizsgálatba beválasztás ideje a kiindulási pont.

az Amerikai Egyesült Államokban az Élelmiszer és Gyógyszerügyi Hivatal (FDA) 1992-ben elfogadta a felgyorsított gyógyszer-befogadási szabályozást, a gyártók gyakran használtak helyettesítő végpontokat, hogy meggyorsítsák az olyan új szerek piacra kerülését, melyek potenciálisan megmenthetnek vagy megnyújthatnak életeket. Ezek a helyettesítő végpontok az onkológi-

ai gyógyszerek befogadásához is használhatóak. Az onkológiában használt helyettesítő végpontok közé tartozik a teljes válaszaránya (ORR), a progressziómentes túlélés (PFS), a betegségmentes túlélés (DFS), a progresszióig eltelt idő (TTP) és a betegetől nyert végpontok, melyeket vagy egyedileg vagy valamilyen összesített végpont részeként közölnek. A helyettesítő végpontok

nagy többsége egy időintervallumot jelent, amely egy specifikus esemény bekövetkeztéig tart. Annak ellenére, hogy néhány helyettesítő végpont hasonlóknak tűnik, nagyon fontos különbségek vannak közöttük, melyeket észben kell tartani, amikor olyan klinikai vizsgálatokat értékelünk, melyekben ezeket a végpontokat használják.

A teljes túlélés továbbra is az arany standard, de egyre inkább úgy gondolják, hogy colorectalis, illetve előrehaladott emlőrákban egyre kevésbé megfelelő mérőszám. Az előrehaladott emlőrákban publikált 58 randomizált klinikai vizsgálatban a teljes túlélés (OS) csak egyszer töltötte be az elsődleges végpont szerepét, míg az időfüggő helyettesítő végpontokat 36 esetben használták. Ez is azt a trendet mutatja, hogy a teljes túlélés helyett helyettesítő végpontokat használnak. A teljes túléléssel összehasonlítva a helyettesítő végpontoknak számos előnye van, de a legfontosabb ezek közül is az, hogy lerövidítik az időt, hogy a betegek a lehető leggyorsabban tudják használni az új kezeléseket. Nagyon fontos, hogy a helyettesítő végpontokat validálják, azaz bebizonyosodjon, hogy adekvátnan helyettesítik a teljes túlélést. E nélkül az újonnan elfogadott gyógyszer akár ártalmas is lehet a betegeknek.

Egy valid helyettesítő végpontnak három statisztikai kívánalmat kell kielégíteni: legyen összefüggésben a daganattal, a kezelés legyen összefüggésben a helyettesítő végponttal, és a helyettesítő végpontnak meg kell teremteni a kapcsolatot a kezelés és a daganat között. Ha ezek az előfeltételek teljesülnek, egy valid helyettesítő végpont megfelelően és precízen előrejelzi a kezelés hatását az igazi végpontra.

Nagyon oda kell figyelni, ha a helyettesítő végpontokat használó klinikai vizsgálatokat hasonlítunk össze, mert egy definiált kezelési modalitásra validált helyettesítő végpont nem biztos, hogy validált egy másik kezelési eljárásra vagy másik tumortípusra. Emiatt minden helyettesítő végpontot minden egyes kezelési szituációra validálni kell. Ennek oka az, hogy egy beavatkozás a helyettesítő végpont által monitorozott más mechanizmuson keresztül is hathat a daganatra. Hasonló megfontolások érvényesek a toxikus hatásokra, melyeket a helyettesítő végpont nem biztos, hogy kimutat.

Általános hibák a klinikai vizsgálatok tervezésében, értékelésében és közlésében

Szelekciós torzítás

Akkor lép fel, ha egy klinikai vizsgálatban szisztematikus különbség van a kezelési karokba történő beválasztásban vagy visszautasításban. Elkerülése valós randomizálással történik, azaz, a beválasztást nem befolyásolja sem a vizsgáló, sem pedig a vizsgálati személy. Sajnos

a vizsgáló a szűrés során befolyásolhatja az eredményt, mert egy meghatározott kezelési karra allokált, de neki nem tetsző beteget a kizárási kritériumoknak jobban megfelelővé tehet, vagy a betegtájékoztató során a beleegyező nyilatkozat visszavonását is elérheti. A tapasztalatok szerint az ilyen szelekciós hiba az új kezelés hatását akár 40%-kal is megnövelheti, emiatt a sikeres szűrés előtt el kell titkolni a beteg kezelési karra történő allokációját. Sajnos ezt csak a klinikai vizsgálatok 10%-ában közlik.

Adatgyűjtési/értékelési torzítás

Az adatgyűjtési/értékelési hiba akkor következik be, ha a klinikai vizsgálat eredményeit és következtetéseit szisztematikus torzítja annak ismerete, hogy melyik vizsgálati személy milyen kezelést kapott. Ezt a hibát előidézheti a kezelést végző vizsgáló, a kezelést kapó résztvevő, az eredményeket értékelő vagy arról beszámoló személy. Elkerülésére szolgál a vizsgálatok lefolytatásának vak módszere (blinding) vagy a maszkolás (masking). Ezeket olyan hosszú ideig fenn kell tartani, ameddig csak lehetséges. A vak stratégiát a vizsgálat során minimum két periódusban kell alkalmazni, először az aktív adatgyűjtés időszakában, a kezeléseket megadásától a végponti adatok begyűjtéséig, majd másodszor az adatok analízisétől az adatközlésig. Nem vak vizsgálatok esetén a vizsgáló szisztematikus képes az eredményt oly módon befolyásolni, hogy az egyik kar betegei részére jobb szupportációt vagy aktívabb mellékhatás-menedzsmentet alkalmaz, és az új kezelést kapó betegek is hajlamosabbak kedvezőbb szubjektív eredményekről beszámolni. Emiatt a legelterjedtebb és legkívánatosabb vak módszer a kettős vak vizsgálat, ahol sem a vizsgáló, sem a vizsgálati személy nem tudja, hogy éppen melyik kezelést adják, illetve kapják. A vak módszer fontosságát mutatja, hogy nyílt vizsgálatok esetében a kísérletes kar eredményei sokszor jobbak, és a nem vak elrendezés 17%-kal javíthatja a vizsgált kezelés adatait. Sajnos a vizsgálatok nagy részében nem közli azt, hogy miként érték el a kettős vak módszer biztonságos alkalmazását. Az adatgyűjtés befejezése után is előidézhető ez a hiba, mert az adatfeldolgozást végző vizsgálók az új kezelést előnyben részesítő végkimeneteli adatokat szelektálhatnak, vagy figyelmen kívül hagyhatják az új kezelést nem támogató eseményeket. Emiatt nem szabadna a kódokat feltörni az adatfeldolgozás lezárásáig.

Fontos megjegyezni a két torzítás közötti különbséget, az egyik az allokáció el nem titkolásából, a másik pedig a vak módszer alkalmazásának hiányából ered. Az allokáció eltitkolása a szelekciós hibát előzi meg, védve a randomizációs szekvenciát a kezelés megadásától, és ez mindig alkalmazható. A vak módszer

az adatgyűjtési/értékelési hibát küszöböli ki, de nem alkalmazható mindig.

A beleegyező nyilatkozatok visszavonásának, a kieső betegeknek illetve a protokollsértéseknek a nem megfelelő kezeléséből fakadó torzítás

Ideális esetben minden résztvevőnek be kellene fejeznie a vizsgálatot, követnie kellene a protokollt, és a végpontok adatait a megfelelő időben ki kellene tudni nyerni. A valóságban azonban minden vizsgálatban vannak hiányzó adatok a betegek kiesése, a protokoll nem megfelelő követése, a nem megfelelő adatgyűjtés miatt. Az ilyen események nem megfelelő kezelése szintén torzításhoz vezethet. Például a betegek kilépése egy vizsgálatból a gyógyszer mellékhatásai miatt elfedheti a valós mellékhatásprofilot. A hiba elkerülésére két megközelítés létezik, az egyik az „intent-to-treat” analízis, ami azt jelenti, hogy az összes beválasztott, azaz kezelni szándékozott beteg adatát fel kell dolgozni olyan kezelési csoportadatként, ahová a beteg eredetileg tartozott. A másik a „worst case scenario”, ahol az előzetes adatanalízis után az elveszett betegekhez a legjobb adatokat rendelik, ha azok a rosszabb eredményű csoportba tartoztak, illetve a legrosszabb adatokat kapják, ha a jobb eredményű csoportba tartoztak, s az így kapott eredményeket az eredetiekkel összevetik.

A keresztezett elrendezésű vizsgálatok nem megfelelő használatából fakadó torzítás

A keresztezett elrendezésű (cross over) vizsgálatok használata nem megengedett véglegesen gyógyítható vagy letális betegségek esetén. Ez az onkológiai klinikai vizsgálatok esetében azt jelenti, hogy a kar váltásának indikátora csakis a vonalváltás szükségessége (progresszió, tűrhetetlen toxicitás) lehet.

Publikációs torzítás

A publikációs torzítás azt jelenti, hogy a tudományos folyóiratok szívesebben publikálnak pozitív vizsgálatokat, míg a nem szignifikáns eredményeket adó vizsgálatokat nem közlik. Ennek a gyakorlatnak az a súlyos következménye, hogy az I. szintű evidenciákat adó metaanalízisek hamis szignifikáns hatásokat mutathatnak. Ez a fajta publikációs torzítás állhat az gyógyszeripar különféle szereplői által lefolytatott klinikai vizsgálatok között talált szisztematikus eltérések hátterében. Ezt a hibát csak a vizsgálatok kötelező regisztrációjával és az eredmények kötelező publikációjával lehetne orvosolni.

Szponzor indukálta torzítás

Ezt a hibát elsősorban ipar szponzorálta vizsgálatokban lehet felfedezni. A hiba lényege az, hogy a vizsgálat során lehet, hogy az összehasonlító kezelést/gyógyszert

szuboptimálisan alkalmazzák/használgják, vagy az adatgyűjtési pontokat úgy határozzák meg, hogy az a vizsgált kezelést/gyógyszert mutassa jobbnak. Emiatt egy közölt klinikai vizsgálat esetében elsődleges fontosságú, hogy áttanulmányozzuk a vizsgálók finansiális érdekeltségeit közlő nyilatkozatokat, és meghatározzuk a lehetséges szponzort.

Irodalom

1. Sackett DL, Rosenberg W, Gray J et al. Evidence based medicine: what it is and what it isn't. *BMJ* 1996;312(7023):71-72.
2. Kaló Z. Beavatkozással nem járó vizsgálatok. In: Lakner G, Renczes G, Antal J (eds). *Klinikai vizsgálatok kézikönyve*. SpringMed Kiadó 2009. pp. 53-56.
3. 15/2012. (VIII. 22.) EMMI rendelet az emberi alkalmazásra kerülő gyógyszerek farmakovigilanciájáról. *Magyar Közlöny* 2012;109:18509 <http://www.kozlonyok.hu/nkonline/MKPDF/hiteles/mk12109.pdf> Accessed 20 July 2014
4. 15/2012. (VIII. 22.) EMMI rendelet az emberi alkalmazásra kerülő gyógyszerek farmakovigilanciájáról. *Magyar Közlöny* 2012;109:18509 <http://www.kozlonyok.hu/nkonline/MKPDF/hiteles/mk12109.pdf> Accessed 20 July 2014
5. Levels of evidence; Centre for Evidence Based Medicine, Oxford http://www.essentialevidenceplus.com/product/ebm_loe.cfm?show=oxford Accessed 20 July 2014
6. Poonacha TK, Go RS. Level of scientific evidence underlying recommendations arising from the National Comprehensive Cancer Network clinical practice guidelines. *Journal of Clinical Oncology* 2011;29(2):186-191.
7. Gachályi B, Lakner G. Gyógyszerfejlesztés. In: Gachályi B, Lakner G, Borvendég J (eds). *Klinikai farmakológia a gyakorlatban - A humán klinikai gyógyszerfejlesztés módszertana* Springer Kiadó, 2002.
8. Senderowicz AM. Information needed to conduct first-in-human oncology trials in the U.S. A view from a former FDA Medical Reviewer. *Clin Cancer Res* 2010;16:1719-25.
9. Lakner G. Generikus gyógyszerfejlesztés, hasonló biológiai gyógyszerek. In: Lakner G, Renczes G, Antal J (eds). *Klinikai vizsgálatok kézikönyve*. SpringMed Kiadó 2009. pp. 58-63.
10. Guidance for Industry Clinical Pharmacology Data to Support a Demonstration of Biosimilarity to a Reference Product May 2014. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM397017.pdf> Accessed 20 July 2014
11. Guidance for Industry Quality Considerations in Demonstrating Biosimilarity to a Reference Protein Product [http://www.fda.gov/downloads/Drugs/February 2012 GuidanceComplianceRegulatoryInformation/Guidances/UCM291134.pdf](http://www.fda.gov/downloads/Drugs/February%202012/GuidanceComplianceRegulatoryInformation/Guidances/UCM291134.pdf) Accessed. 20 July 2014.
12. Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: non-clinical and clinical issues http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/06/WC500144124.pdf Accessed on 20 July 2014. július 27.
13. Guideline on immunogenicity assessment of biotechnology-derived therapeutic proteins April 2008 http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003946.pdf Accessed: 20 July 2014
14. Lakner G, Gachályi B, Singer J [eds]. *Klinikai farmakológia – Biostatistikai fogalomtárral*. Springmed Kiadó, Budapest; 2005. pp. 320-1.
15. Guidance for industry: Clinical trial endpoints for the approval of cancer drugs and biologics. Food and Drug Administration, May 2007) <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071590.pdf> Accessed 20 July 2014