HODNOCENÍ KLINICKÝCH STUDIÍ, ZÁKLADNÍ POJMY

MUDr. Michaela Matoušková
Urocentrum Praha

Při objektivním hodnocení účinnosti a bezpečnosti léčivých přípravků a zdravotnických prostředků jsou předkládány veřejnosti výsledky multicentrických randomizovaných klinických studií. Design a analýza klinických studií vychází ze znalostí medicíny a farmakologie, ale i statistiky, matematiky a informatiky. V současnosti problematika využívá již vlastních postupů a terminologie. Cílem sdělení je zopakovat základní pojmy, se kterými se v klinické praxi při analýze dat setkáváme.

Urolog. pro Praxi, 2007; 1: 37–38

Objektivní hodnocení účinnosti a bezpečnosti léčivých přípravků a zdravotnických prostředků vychází z dat získaných v rámci randomizovaných klinických studií. Tvorba, sběr dat, design studie a jejich hodnocení jsou dnes samostatnou disciplínou. Výstupy pak obvykle vedou k registraci nových přípravků. Z hlediska vývoje nových léčivých přípravků rozlišujeme období preklinické a klinické. Preklinický vývoj studuje nové molekuly na buněčné úrovni a zvířecích modelech. Vlastní klinické studie hodnotí experimenty ve fázi předregistrační (fáze I–III) a poregistrační (fáze IV). Samostatnou skupinu dále tvoří studie postmarketingové (neintervenční) a specifické léčebné programy.

Studie fáze I

Hlavním cílem klinických studií fáze I je stanovení základních humánních farmakokinetických parametrů nových léčivých přípravků a stanovení maximální tolerovatelné dávky (MTD). Důraz je tedy kladen primárně na bezpečnost léčby. Studie fáze I navazují bezprostředně na experimenty prováděné na zvířecích modelech a sledují především nežádoucí účinky. Do klinických hodnocení fáze I je zařazován relativně malý počet subjektů, zpravidla 12–20, obvykle zdravých dobrovolníků. K testování toxických léčivých přípravků (cytostatika apod.) jsou zařazováni pacienti. Uspořádání experimentů umožňuje matematické modelování vztahu „dávka–odpověď" a odhad efektu různých koncentrací látky na lidský organizmus.

Studie fáze II

Klinické studie fáze II navazují na vyhodnocené studie fáze I. Hlavním cílem studií fáze II je zjištění skutečné účinnosti hodnoceného přípravku ke zdůvodnění jeho dalšího testování v experimentech pokročilejší fáze. Souběžným cílem studií fáze II je rozšířené hodnocení tolerance a bezpečnosti sledované léčby. Do těchto experimentů bývá zpravidla zařazováno 20–200 subjektů. Většinou se jedná o jednoramenné experimenty, kde je zjištěná účinnost a bezpečnost srovnávána s literárně dostupnými údaji o srovnatelných přípravcích nebo o současně nejlepší dostupné a používané léčbě.

Studie fáze III

Klinické studie fáze III představují nejznámější a nejcharakterističtější typ klinického hodnocení. Jedná se o sledování, při kterém je nejčastěji srovnávána účinnost a bezpečnost dvou a více léčebných postupů v rámci jednoho randomizovaného sledování. Cílem těchto studií je tedy srovnat vlastnosti nového léčivého přípravku s kontrolou, kterou může představovat placebo nebo alternativní léčebný postup či léčivý přípravek. Závěry těchto studií jsou podkladem pro schválení přípravku, umožňující jejich klinické podávání (např. FDA, SÚKL a další). Sledování zahrnuje sta až tisíce subjektů, ale v závislosti na designu studie a epidemiologii onemocnění se mohou počty zásadně lišit. Nejčastěji bývají studie fáze III uspořádány podle paralelního či cross-over designu. Při klasickém paralelním uspořádání jsou pacienti randomizováni do dvou a více ramen. Randomizace přiřadí jednotlivé subjekty hodnocení do srovnávaných ramen studie a v těchto ramenech, reprezentujících vždy jeden jediný léčebný postup, zůstávají subjekty hodnocení v průběhu celého experimentu. Při cross-over uspořádání jsou subjekty hodnocení také randomizovány do jednotlivých ramen studie. Tato ramena však představují pouze pořadí, ve kterém pacienti podstoupí dva nebo více léčebných schémat srovnávaných v daném experimentu. Vzhledem k tomu, že při tomto uspořádání je každý subjekt „sám sobě kontrolou", je z hlediska statistické analýzy výhodnější a vyžaduje nižší počet subjektů ve srovnání s paralelním uspořádáním.

Studie fáze IV

Klinické studie fáze IV jsou prováděny v období po úspěšné registraci nových léčivých přípravků, tedy v době jejich běžného používání v klinické praxi. Hlavním cílem těchto projektů je potvrzení vlastností hodnocených přípravků známých z předchozích experimentů (ze studií fáze I–III) na širších populacích pacientů (bez výraznějšího omezení přísnými vstupním a vylučujícími kritérii) a za „reálných" podmínek klinické praxe. Tyto studie vzhledem k vysokému počtu zařazených pacientů bývají často zaměřeny na detailní analýzu nežádoucích účinků léčby, především vzácnějších projevů, které nemusely být zachyceny v předchozích studiích.

Zaslepení

Při zaslepení nejsou schopni ani hodnocený, ani ostatní účastníci na základě vlastností podávaného přípravku (vzhled, chuť, velikost, způsob aplikace apod.) rozlišit, který z přípravků je podáván. Cílem je zvýšení objektivosti hodnocení účinnosti a bezpečnosti a zvýšení validity výsledků. Při jednoduchém zaslepení je zalepena pouze jedna strana, buď subjekt hodnocení nebo investigátor, při zaslepení obou hovoříme od „double blinding" (dvojitém zaslepení). Pokud je zaslepen i personál spravující data, jedná se o trojité zaslepení (triple blinding), termín čtyřnásobné zaslepení (quadruple blinding) se používá tehdy, jestliže jsou osoby aplikující studijní medikaci odlišné od osob vyhodnocujících léčebný efekt a bezpečnost a tyto osoby jsou rovněž zaslepeny. Někdy je zaslepení nerealizovatelné, např. při srovnání medikamentózní a chirurgické léčby.

Randomizace

Randomizací rozumíme proces náhodného rozdělování hodnocených subjektů do dvou či více léčebných skupin srovnávaných v rámci klinického hodnocení. Zlatým standardem je použití randomizace u srovnávacích klinických studií fáze III s cílem zamezit selektivnímu zařazování konkrétních pacientů do konkrétních ramen studie. Navíc moderní postupy randomizace umožňují rovnoměrné rozdělení hodnocených ve srovnávaných ramenech i v čase (umožňují provedení interim analýz) a zajišťují rovnoměrnou distribuci určených prognostických faktorů.

Randomizační techniky, se kterými se v současné době setkáváme, rozdělujeme z hlediska vhodnosti na nepřípustné, méně vhodné a doporučené. Mezi nepřípustné randomizační techniky patří v minulosti často používané randomizace subjektů hodnocení

na základě jejich pořadového čísla vstupu do studie (např. každý „lichý“ pacient zařazen do ramene A, každý „sudý“ do ramene B) a randomizace podle iniciál pacienta (např. pacienti se jmény začínajícími písmenem z první poloviny abecedy zařazeni do ramene A, ostatní do ramene B). Analogicky nepřípustná je randomizace na základě data narození nebo data vstupu do studie. Nevhodnost těchto technik je dána především tím, že v případě otevřeného klinického hodnocení (pacient ani zkoušející lékař nejsou zaslepeni) může investigátor ovlivnit zařazení konkrétního subjektu do konkrétní léčebné skupiny tím, že ještě před podpisem informovaného souhlasu ví, do které léčebné skupiny by byl pacient zařazen, a účast v daném experimentu nemusí na základě této znalosti pacientovi vůbec nabídnout. U zaslepených studií je nevýhodou těchto postupů nemožnost kontrolovat distribuci prognostických faktorů v ramenech studie a nemožnost zajistit požadovaný poměr počtu subjektů v léčebných skupinách. Za méně vhodné randomizační techniky je považována např. tzv. kompletní randomizace, tedy technika, kdy po podpisu informovaného souhlasu pacienta je o jeho přiřazení do konkrétního ramene studie rozhodnuto pouze na základě hodu mincí (nebo tahu losu, použití náhodného generátoru čísel apod.). Tato technika je považována za méně vhodnou, protože sice znemožňuje „předvídat“ experimentátorem přiřazení konkrétního subjektu konkrétní léčebné skupině, na druhé straně však opět není možno kontrolovat dosažení požadovaného poměru počtu subjektů v ramenech studie v jejím průběhu a už vůbec není kontrolována distribuce prognostických faktorů. K doporučovaným randomizačním technikám patří především stratifikovaná permutační bloková randomizace a adaptivní randomizace. Stratifikovaná randomizace vychází z principu, kdy jsou prospektivně definovány nejdůležitější prognostické faktory, které mohou u konkrétního subjektu zásadně ovlivnit účinnost a bezpečnost hodnoceného léčebného režimu (např. stadium onemocnění, věk, pohlaví aj.) a jsou popsány všechny teoreticky možné kombinace těchto faktorů u konkrétních jedinců, tzv. strata (tedy jedno stratum mohou např. představovat ženy mladší 45 let s klinickým stadiem 1). Randomizace je následně provedena samostatně v rámci všech těchto strat, a je tak zajištěno, že ve všech ramenech jsou rovnoměrně zastoupeni pacienti s časným a pokročilým stadiem, že bude mezi rameny stejný podíl mužů a žen a že mezi rameny budou srovnatelně zastoupeny všechny věkové skupiny. Pro zajištění požadovaného poměru počtu subjektů v ramenech studie v jejím průběhu je v případě permutační blokové randomizace použito náhodně opakované sekvence pořadí jednotlivých ramen studie, v důsledku čehož je vždy po daném počtu subjektů (např. při velikosti bloku n = 4) po každých čtyřech subjektech hodnocení zajištěn vyrovnaný počet pacientů ve všech ramenech. V případě adaptivních randomizačních technik se narozdíl od výše uvedených postupů nevytváří prospektivní randomizační plán, ale randomizace každého nově zařazeného subjektu hodnocení se provádí až po vyhodnocení dosavadního průběhu experimentu, a to buď z hlediska průběžně vyhodnocené účinnosti léčby, nebo distribuce prognostických faktorů.

Velikost souboru

Počet subjektů hodnocení zařazených do řádně designovaného klinického hodnocení není náhodný a je přesně stanoven statistikem na základě tzv. „power analýzy“. Při zařazení menšího počtu sledovaných je pravděpodobné, že výsledek studie nebude mít dostatečnou „sílu“ prokázat hypotézu, např. vyšší účinnost a bezpečnost nového přípravku, a to i přesto, že je nový přípravek bezpečnější a efektivnější. Ani zbytečně velké množství sledovaných není přínosné. Vždy je nezbytné optimalizovat velikost zkoumaného vzorku populace. Cílem optimalizace je zařadit do sledování takový počet subjektů, který bude dostatečný ke zjištění statisticky významného rozdílu v účinnosti a/nebo bezpečnosti mezi rameny studie. Tzn. že pravděpodobnost dosažení výsledku náhodně je minimální. Důvody nezbytnosti optimalizace velikosti vzorku jsou etické, ekonomické i legislativní. Je neetické vystavovat působení experimentálního přípravku více osob, než je nezbytné. Nadbytečné počty sledování zvyšují neúměrně ekonomické náklady. A konečně metodika výpočtu velikosti vzorku je součástí protokolu studie, který je před zahájením klinického sledování dokládán institucím, jako je FDA či SÚKL.

Interim analýza

Jde o jednorázové nebo opakované testování primární hypotézy v průběhu sledování před plánovaným ukončením studie. Cílem interim analýzy je uvážení možnosti předčasného ukončení studie v případě, že existuje vážný předpoklad, že další pokračování experimentu již nepřinese žádnou nosnou informaci. Hlavním cílem co nejčasnějšího ukončení klinického hodnocení je urychlit přístup pacientů k nové léčbě, nebo naopak minimalizace počtu pacientů vystavených léčbě, která se neprokázala jako bezpečná. Nezanedbatelná je rovněž úspora finančních nákladů a úspora lidské práce související s pokračováním studie. Již v protokolu studie musí být uvedeno kdy a kolik interim analýz má být provedeno, zároveň je nezbytné modifikovat hladinu významnosti právě podle počtu plánovaných interim analýz a nastavit si tak „přísnější“ kritéria pro průkaz lepší účinnosti či bezpečnosti hodnoceného přípravku při opakovaném testování primární hypotézy.

Metaanalýza

Výsledky jednotlivých studií s konkrétními přípravky či postupy přinášejí často výstupy odlišné, mnohdy protichůdné, ačkoliv design studie, indikace a populace pacientů byly shodné. Zobecnění výsledků dvou a více nezávislých studií přináší metaanalýza. Metodika je založena na definování vstupních a vylučujících kritérií pro jednotlivé klinické studie a na aplikaci speciálních statistických technik, jejichž cílem je analýza heterogenity výsledků jednotlivých experimentů. V případě průkazu heterogenity je tato interpretována a při potvrzení homogenit jsou výsledky zobecněny. Cílem metaanalýz je přinést ucelenou a klinicky relevantní informaci na základě výsledků jednotlivých klinických studií, které mohou mít nejednotné závěry.

Při zpracování metaanalýz se statistik potýká s různými možnostmi zkreslení. Jedním ze základních okruhů je problém tzv. „publication bias“. Ke zkresleným závěrům může totiž dojít, pokud jsou pro mataanalýzu použity pouze publikované výsledky a nejsou zahrnuta nepublikovaná data. Je prokázáno, že studie s relativně malým počtem hodnocených subjektů a negativním výsledkem (když např. výsledek neprokázal přínosnost nového farmaka) mají významně menší šanci být publikovány v odborném tisku, přestože prezentují pravdivou informaci. Jinou rovinou je nejednotnost hodnocení účinnosti a bezpečnosti v jednotlivých studiích, např. doba do progrese je v jednotlivých studiích hodnocena různě, jednou od randomizace, jindy od doby stanovení diagnózy či od zahájení léčby atd. Důležitým problémem jsou vícenásobné publikace výsledků studie na identické populaci subjektů. To je důsledkem logického procesu, kdy autoři po ukončení sledování mají tendenci k co největšímu počtu publikací výsledků, často s rozdílnými názvy a koncipovanými tak, aby nebylo snadné zjistit, že se jedná o shodný projekt a neměnnou populaci nemocných.

Závěrem

Pokud jste vydrželi s četbou až do tohoto místa, pak byste měli mít alespoň hrubou orientaci o tom, jak a proč jsou koncipovány klinické studie a jak se dívat na jejich závěry. Medicínská statistika se v posledních desetiletích vyvinula ve speciální odvětví. S jejími výstupy pracujeme v klinické praxi každodenně. Každodenně tak stojí při našem rozhodování o zahájení či pokračování léčby na jedné straně výsledky studií předkládané firmami a na druhé naše osobní zkušenost.

MUDr. Michaela Matoušková
Urocentrum Praha
Karlovo nám. 3, 120 00 Praha 2
e-mail: matouskova@urocentrum.cz