

TEORIE A PRAXE KLINICKÝCH HODNOCENÍ V ONKOLOGII

CLINICAL EVALUATION IN THEORY AND PRACTICE IN ONCOLOGY

klinická onkologie

Supplement 2005

ROČNÍK 18

30. prosince 2005

VYDÁVÁ
ČESKÁ LÉKAŘSKÁ
SPOLEČNOST J. E. PURKYNĚ
IČO 444359

V NAKLADATELSTVÍ
ApS BRNO, spol. s r. o.
IČO 543535

REDAKCE:
Masarykův onkologický ústav Brno
Žlutý kopec č. 7
656 53 Brno

Sekretář redakce:
ing. Zdeněk Bouša
tel., fax: 543 134 226

Grafická a technická úprava:
Bohuslav Havlíček

Tiskne Moravská typografie, a. s.
Brno, Moravské náměstí 13
IČO 15549763

Vychází 6krát ročně
Roční předplatné 180 Kč
pro studenty LF 90 Kč

Expedici na základě roční objednávky
vyřizuje redakce

Ministerstvo kultury ČR
MK ČR 5158
ISSN 0862-495 X

INTERNET – vstupní adresa:
<http://www.linkos.cz>

INDEXED IN EXCERPTA MEDICA

ČASOPIS ČESKÉ ONKOLOGICKÉ SPOLEČNOSTI A SLOVENSKEJ ONKOLOGICKEJ SPOLEČNOSTI

THE JOURNAL OF THE CZECH AND SLOVAK ONCOLOGICAL SOCIETIES

VEDOUcí REDAKTOR: REJTHAR ALEŠ

**ZÁSTUPCE VEDOUcíHO
REDAKTORA:** KOZA IVAN
VÝKONNÝ REDAKTOR: FAIT VUK

REDAKTOŘI:

ČOUPEK PETR
HÁJEK ROMAN
KOCÁK IVO

VALÍK DALIBOR
VORLÍČEK JIŘÍ
ŽALOUĐÍK JAN

REDAKČNÍ RADA:

ADAM ZDENĚK, Brno
BABUŠÍKOVÁ OLGA, Bratislava
BEDNAŘÍK OTAKAR, Brno
BILDER JOSEF, Brno
ČOUPEK PETR, Brno
DRBAL JOSEF, Brno
ECKHARDT SANDOR, Budapešť
FAIT VUK, Brno
HÁJEK ROMAN, Brno
JURGA LUDOVIT, Trnava
KALLAY JOZEF, Bratislava
KAUŠITZ JURAJ, Bratislava
KLAŠTERSKÝ JAN, Brusel
KLENER PAVEL, Praha
KOCÁK IVO, Brno
KOUTECKÝ JOSEF, Praha

KOVAŘÍK JAN, Brno
KOZA IVAN, Bratislava
MAYER JIŘÍ, Brno
MECHL ZDENĚK, Brno
NĚMEC JAROSLAV, Brno
ONDRUŠ DALIBOR, Bratislava
PAČOVSKÝ ZDENĚK, Brno
PLEŠKO IVAN, Bratislava
PETRUŽELKA LUBOŠ, Praha
REJTHAR ALEŠ, Brno
SPURNÝ VLADIMÍR, Brno
UJHÁZY VILIAM, Bratislava
VORLÍČEK JIŘÍ, Brno
VYZULA ROSTISLAV, Brno
WAGNEROVÁ MÁRIA, Košice
ŽALOUĐÍK JAN, Brno

TEORIE A PRAXE KLINICKÝCH HODNOCENÍ V ONKOLOGII

OBSAH

2005

Úvodní slovo Žaloudík J., Vyzula R.	237
Svobodník A., Coufal O., Dušek L. Základní pojmy v designu, analýze dat a interpretaci výsledků klinických hodnocení léčiv	238
Demlová R. Základní typy parametrů hodnocených v klinických studiích v onkologii	242
Tomášek J., Kiss I. Design klinické studie	245
Demlová R. Hodnocení odpovědi nádoru na léčbu - recist (Response Evaluation Criteria In Solid Tumors)	248
Demlová R. Výběr pacientů v klinických studiích - rizika pro interpretaci v běžné klinické praxi	250
Novotný J., Petruželka L. Metodické problémy klinických studií v onkologii	252
Žaloudík J., Vyzula R. Odrz výsledků klinických studií v populačním onkologickém registru	256

CLINICAL EVALUATION IN THEORY AND PRACTICE IN ONCOLOGY

CONTENTS

2005

Introduction Žaloudík J., Vyzula R.	237
Svobodník A., Coufal O., Dušek L. Design, data analyses and interpretation of clinical trials-basic concepts	238
Demlová R. Primary end-points in clinical trials in oncology	242
Tomášek J., Kiss I. Design of a clinical trial	245
Demlová R. Selection of patients in clinical trials - risks for interpretation in routine practice	248
Demlová R. Response evaluation criteria in solid tumors	250
Novotný J., Petruželka L. Methodological problems of clinical trials in oncology	252
Žaloudík J., Vyzula R. Reflection of results from clinical trials in population cancer registry	256

ÚVODNÍ SLOVO

ŽALOUĐÍK J., VYZULA R.

MASARYKŮV ONKOLOGICKÝ ÚSTAV V BRNĚ

Časopis Klinická onkologie předkládá odborné onkologické veřejnosti formou suplementa první ucelený soubor prací o metodice klinických studií, respektive klinických hodnocení jak zní správnější označení. Kontrolované klinické studii jsou nesporně základním nástrojem vědeckého bádání v klinické praxi a hledání optimálních léčebných postupů. Všichni z nich v terapeutickém rozhodování ve své praxi vycházíme. Nejsme si však zcela jisti, že jejich metodologii rozumíme všichni stejně a že si rozumíme i navzájem při jejich interpretacích.

Abychom pomohli tyto obavy rozptýlit, vznikl soubor prací, k němuž bude možné se i v českém onkologickém písemnictví odvolat jako k základnímu a referenčnímu metodologickému zdroji. Nenutili jsme autory jednotlivých článků, aby se zcela vyhnuli anglické terminologii, která v metodologii klinických hodnocení převládá. Není totiž nezbytné se snažit za každou cenu o nešikovné české překlady za situace, kdy pojednáváme o klinických studích, které jsou organizovány v drtivé většině v zahraničí a pro něž je anglická terminologie normou.

Kromě základních poznatků byl v suplementu dán prostor také názorům, které různé aspekty kontrolovaných klinických hodnocení poněkud problematizují nebo konfrontují s běžnou praxí. Má to přispět ke kritickému myšlení našich onkologů, pro které musejí být závěry klinických studií sice i nádále hlavním vodítkem, nikoli však dogmatem. Současný důraz na standardizaci medicíny má směřovat ke standardizaci a sjednocování nástrojů, nikoli bránit racionální individualizaci léčebných postupů ve prospěch nemocného.

Léčebná péče se také v onkologii odvíjí od stálé interakce s nabídkou a lékovým vývojem velkých farmaceutických firem. Vznik suplementa podpořila firma Roche. Nesporně jí patří dík za to, že vnímá jak cenným spolupracovníkem jsou onkologové nejen informovaní, ale také dostatečně kritičtí, aby medicína nepřestávala sloužit nemocným a nové poznatky byly do praxe uváděny na vědeckém základě.

Věříme, že suplementum pomůže v orientaci zejména mladým lékařům, kteří hledají pro svou praxi hlavní záchytné informace a vzorce pro aplikaci medicíny založené na průkaznosti. Přejeme všem našim lékařům i studentům, aby toto suplementum shledali užitečným příspěvkem ke zdrojům českého onkologického písemnictví.

ZÁKLADNÍ POJMY V DESIGNU, ANALÝZE DAT A INTERPRETACI VÝSLEDKŮ KLINICKÝCH HODNOCENÍ LÉČIV

DESIGN, DATA ANALYSES AND INTERPRETATION OF CLINICAL TRIALS-BASIC CONCEPTS

SVOBODNÍK A., COUFAL O., DUŠEK L.

CENTRUM BIostatistiky a ANALÝZ LF A PŘF MASARYKOVY UNIVERZITY V BRNĚ

Souhrn

Metodika designu a analýzy dat randomizovaných klinických studií představuje v současnosti samostatnou vědní disciplínu s vlastní terminologií. Vzhledem k tomu, že nauka o provádění a hodnocení výsledků klinických studií není součástí standardního vzdělávání lékařů, může docházet k desinterpretaci publikovaných dat. V textu jsou diskutovány základní pojmy z experimentálního designu a analýzy klinických studií.

Klíčová slova: klinické studie, randomizace, analýza dat, interim analýza, metaanalýza

Summary

Methodology of the design and analysis of clinical trials currently represents a substantive branch of knowledge with its own terminology. As courses on this subject are not incorporated in the standard educational process in medicine, misinterpretation of published results of clinical trials could occur. In this paper are discussed basic terms used in the design and analysis of randomized clinical trials.

Key words: clinical trials, randomization, data analysis, interim analysis, metaanalysis

ÚVOD

Randomizovaná klinická hodnocení (randomizované klinické studie) představují v současnosti nejlepší dostupnou technologii pro objektivní zhodnocení vlastností (především účinnosti a bezpečnosti) léčivých přípravků a zdravotnických prostředků. Metodika a zásady správného provádění klinických studií jsou dnes již podrobně zpracovány, definovány a standardizovány. Rychlý vývoj této oblasti v posledních desetiletích vedl k tomu, že se problematika designu a analýzy dat klinických studií poněkud vzdálila příbuzným oborům (medicína, farmakologie, matematika, statistika, informatika aj.) a komunikuje dnes vlastní terminologií a jazykem. Vzhledem k tomu, že nauka o designu a analýze dat klinických studií není součástí standardního vzdělávacího procesu uživatelů výsledků těchto experimentů (tedy především lékařů), může snadno docházet k desinterpretaci publikovaných dat.

Cílem tohoto textu je seznámit čtenáře s vybranými základními pojmy designu klinických studií. Tento text neumí nahradit úplné vzdělání v dané problematice a všem zájemcům o podrobnější informace je doporučeno samostudium (existuje velké množství kvalitní a dostupné literatury) nebo návštěva specializovaných vzdělávacích kurzů u nás nebo v zahraničí.

ZÁKLADNÍ KLASIFIKACE KLINICKÝCH STUDIÍ

Existuje několik přístupů ke klasifikaci klinických studií, nejpoužívanější je třídění z hlediska pozice v procesu vývoje a registrace nových léčivých přípravků. Podle této klasifikace se experimenty dělí na prováděné v období *preklinického vývoje* (studie na zvířecích modelech) a na experimenty prováděné v období *klinického vývoje* (vlastní klinické studie), které se dále dělí na studie 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). Vzhledem k tomu, že tyto studie navazují bezprostředně na experimenty prováděné na zvířecích modelech, jsou v těchto projektech podrobně sledovány a analyzovány nežádoucí účinky přípravků, důraz je tedy kladen primárně na bezpečnost léčby.

Do klinických hodnocení Fáze I je zařazován relativně malý počet subjektů hodnocení (zpravidla 12-20), většinou se jedná o zdravé dobrovolníky, výjimkou je testování toxických léčivých přípravků (cytostatika apod.), v těchto případech jsou zařazováni pacienti. Tyto experimenty jsou zpravidla uspořádány tak, aby výsledná data umožňovala matematické modelování vztahu „dávka-odpověď“ a odhad efektu různých koncentrací látky na lidský organismus.

Studie Fáze II

Klinické studie Fáze II navazují na provedené a vyhodnocené studie Fáze I a jejich hlavním cílem je zjištění skutečné účinnosti hodnoceného přípravku za účelem opodstatnění jeho dalšího testování v experimentech pokročilejší fáze (studie Fáze III). Dalším cílem studií Fáze II je souběžné rozšířené hodnocení tolerance a bezpečnosti sledované léčby. Do těchto experimentů bývá zpravidla zařazováno 20-200 subjektů hodnocení. Jedná se většinou o jednoramenné experimenty, kde zjištěná účinnost a bezpečnost je srovnávána s literárně dostupnými údaji o srovnatelných přípravcích nebo současně nejlepší dostupné a používané léčbě.

Studie Fáze III

Klinické studie Fáze III představují zřejmě nejnámější a nejcharakterističtější typ klinického hodnocení. Jedná se o experimenty, ve kterých je nejčastěji srovnávána účinnost a bezpečnost dvou a více léčebných postupů v rámci jednoho randomizovaného

experimentu. Cílem těchto studií je tedy srovnat vlastnosti nového léčivého přípravku s kontrolou, kterou může představovat placebo, nebo „aktivní“ kontrola, tedy alternativní léčebný postup nebo léčivý přípravek. Výsledky těchto studií jsou základními podklady pro schválení používání přípravku regulačními autoritami (např. FDA, SÚKL apod.). Do těchto experimentů bývá zpravidla zařazováno 100-1000 subjektů, počty se však mohou značně lišit v závislosti na designu studie a epidemiologii hodnoceného onemocnění, limitace mohou být dány praktickými možnostmi náběru subjektů. Studie Fáze III bývají nejčastěji uspořádány podle *paralelního* nebo *cross-over* designu. Při klasickém paralelním uspořádání platí, že při srovnání dvou a více ramen studie je základním prvkem designu randomizace, která 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. Naproti tomu 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 (tedy ze studií Fáze I-III) na širších populacích pacientů (bez výraznějšího omezení přísnými vstupními a vylučujícími kritérii) a za „reálných“ podmínek klinické praxe (na rozdíl od „skleníkového“ prostředí klinických studií Fáze I-III). 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

O *zaslepení* hovoříme v klinickém hodnocení tehdy, pokud je cílem experimentu srovnání dvou nebo více léčivých přípravků, a kdy v zájmu objektivního zhodnocení jejich účinnosti a bezpečnosti je subjekt hodnocení nebo i další účastníci studie neschopen na základě vlastností podávaného přípravku (způsob aplikace, vzhled, chuť, velikost apod.) rozlišit, který ze srovnávaných přípravků je aplikován. O *jednoduchém zaslepení (single blinding)* hovoříme tehdy, jestliže je zaslepena pouze jedna strana a to buď subjekt hodnocení nebo investigátor, *dvojitě zaslepení (double blinding)* se používá tehdy, jestliže jsou zaslepeny subjekty hodnocení i investigátoři, termín *trojitě zaslepení (triple blinding)* se používá v případech, že je kromě subjektů hodnocení, a investigátorů zaslepen rovněž personál spravující data studie a konečně 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. Zaslepení považujeme za další prvek v designu klinického hodnocení zvyšující objektivnost a validitu výsledků, v některých případech je však samozřejmě nerealizovatelné (např. srovnání chirurgické léčby s jen medikamentózní).

RANDOMIZACE

Randomizací rozumíme proces náhodného (nebo pseudonáhodného) rozdělování subjektů hodnocení do dvou nebo více léčebných skupin srovnávaných v rámci klinického hodnocení. Použití randomizace je dnes zlatým standardem především při provádění srovnávacích klinických studií Fáze III, tedy v projektech bezprostředně předcházejících registraci nového léčivého přípravku. Hlavním cílem randomizace je zamezit subjektivnímu a selektivnímu zařazování konkrétních subjektů hodnocení do konkrétních léčebných skupin (ramen studie) zkoušejícím lékařem. Moderní mechanismy randomizace umožňují navíc zajistit rovnoměrný

(nebo jiný požadovaný) poměr počtu subjektů hodnocení ve srovnávaných ramenech i v průběhu studie (možnost provedení interim analýz) a zajistit rovnoměrnou distribuci vybraných prognostických faktorů. I přes zjevný přínos randomizace k validitě a objektivitě výsledků srovnávacích experimentů má tato technika i v současnosti některé odpůrce (etické důvody) a první řádně provedené randomizované klinické studie byly provedeny teprve po druhé světové válce, tedy z hlediska celkového vývoje metodiky v relativně nedávné době.

V současnosti nejčastěji používané randomizační techniky můžeme rozdělit podle vhodnosti jejich použití na skupinu technik *nepřípustných, méně vhodných a doporučených*.

Mezi *nepřípustné* randomizační techniky tak patří v minulosti často používané randomizace subjektů hodnocení na základě jejich *pořadového čísla* vstupu do studie (např. každý „lichý“ pacient zařazen do ramene A a každý „sudý“ pacient do ramene B), dále randomizace podle *iniciál pacienta* (např. pacienti se jmény začínajícími písmenem z první poloviny abecedy zařazení do ramene A a ostatní do ramene B), analogicky nepřijatelná 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 experimentátor 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. V případech zaslepených studií je nevhodou 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, kterou si lze zjednodušeně představit tak, že po podpisu informovaného souhlasu pacienta je o jeho přiřazení do konkrétního ramene studie rozhodnuto pouze na základě hodů 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. Mezi *doporučované* randomizační techniky dnes 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 muži mladší 45 let s klinickým stadiem I). 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 na rozdíl 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.

Požadavek nutnosti zajištění kontroly distribuce prognostických faktorů v ramenech studie dokládají příklady následujících dvou onkologických studií. Ve studii autorů Twelves et al. 2005, ve kte-

ré byla srovnávána účinnost a bezpečnost léčby fluorouracilem a leucovorinem versus orálně podávaného fluoropyrimidin capcitabinu u pacientů s metastatickým kolorektálním karcinomem bylo prokázáno, že z pohledu DFS (disease free survival) a OS (overall survival) je léčba capcitabinem minimálně stejně účinná jako léčba fluorouracilem a leucovorinem. V této studii byly nicméně zjištěny rozdíly v hodnotách významných prognostických faktorů (hladina karcinoembryonálního antigenu a stupeň zasažení lymfatických uzlin) v ramenech studie a po adjustaci na tyto rozdíly byla prokázána vyšší účinnost capcitabinu z hlediska DFS i OS. Naproti tomu např. ve studii autorů Hurwitz et al. 2004, srovnávající účinnost a bezpečnost léčby irinotecanem, fluorouracilem, leucovorinem (IFL) a bevacizumabem ve srovnání s IFL a placebem byla randomizace pacientů stratifikována podle centra, vstupních hodnot ECOG performance status, lokalizace primárního nádoru a počtu metastaticky postižených míst. Díky tomuto postupu bylo dosaženo srovnatelného zastoupení pacientů s různou prognózou v ramenech studie a byla validně prokázána vyšší účinnost léčby při použití bevacizumabu.

INTERIM ANALÝZA

Interim analýzou v klinickém hodnocení léčiv rozumíme jednorázové nebo opakované testování primární hypotézy (tedy např. zda je experimentální léčba účinnější než kontrolní) před protokolem plánovaným řádným ukončením experimentu. Cílem interim analýzy je zvážit možnost předčasného ukončení studie v případě, že existuje vážný předpoklad, že další pokračování experimentu již nepřinese žádnou novou informaci. Hlavním cílem co nejdříve 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.

Z výše uvedených důvodů se zdá být logické provádět interim analýzy co nejčastěji a neomezovat se v počtu testování výsledků studie v jejím průběhu. Na druhé straně je však důležité si uvědomit, že opakovaným testováním primární hypotézy se zvyšuje pravděpodobnost průkazu falešně pozitivního výsledku, tedy potvrzení určité konkrétní léčby jako účinnější a bezpečnější i v případě, že tato léčba ve skutečnosti účinnější a bezpečnější není. Abychom si dokázali lépe představit tento problém, demonstrováme ho jednoduchou analogií. Představme si, že budeme v rámci experimentu srovnávat účinnost dvou ve skutečnosti naprosto stejně účinných léčivých přípravků a naším primárním parametrem účinnosti bude podíl pacientů s dosaženou pozitivní léčebnou odpovědí. Tato situace je podobná příkladu, kdy si budeme házet vyváženou mincí, a vždy po určitém počtu hodů vyhodnotíme počet padnutí „lícové“ strany mince (reprezentující počet pacientů s dosaženou léčebnou odpovědí v jednom rameni) a počet padnutí „rubové“ strany mince (reprezentující počet pacientů s dosaženou léčebnou odpovědí ve druhém rameni studie). Představme si, že máme jako vlastníci patentu na syntézu určité molekuly zájem na tom, aby jako účinnější léčba bylo prokázáno rameno, které u nás představuje „lícovou“ stranu mince. Pokud bychom mohli kdykoliv v průběhu studie (pro nás analogicky po jakémkoliv počtu hodů mincí) ukončit experiment, házeli bychom mincí tak dlouho, dokud by se situace nevyvíjela v náš prospěch (tedy dokud by nepřevážil počet hodů, kdy padnul „líc“ nad počtem hodů, kdy padnul „rub“) a v tento okamžik bychom experiment ukončili a prohlásili, že naše léčba byla prokázána jako účinnější a bezpečnější. O tom, že v průběhu testování (házení mincí) nastanou opakovaně a čistě náhodou situace, kdy jednoznačně převažuje počet padnutí „lícové“ nebo „rubové“ strany mince se můžeme snadno přesvědčit (při dostatečné trpělivosti) opakovaním dostatečného počtu hodů (solistickou nebo je možné ověřit tento fakt počítačovou simulací třeba i na datech přežití).

Z důvodu výše uvedených protichůdných požadavků na průběžné hodnocení výsledků klinických studií (jednoznačné *pro* z důvodu možnosti ukončit studii pokud další průběh je již zbytečný a jednoznačné *proti* z důvodu možnosti manipulace

s výsledky) jsou v současnosti interim analýzy vysoce doporučovány, ale pouze při respektování určitých jednoznačně definovaných pravidel. Mezi tato pravidla patří např. povinnost plánovat každou interim analýzu ještě před vlastním zahájením experimentu, tedy již v protokolu studie. V tomto dokumentu musí být jednoznačně uvedeno kdy (tedy např. po jakém přesném počtu zařazených pacientů) bude interim analýza provedena a kolik je jich v průběhu celé studie plánováno. Zároveň je nutné modifikovat hladinu významnosti a 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

Zřejmě každý, kdo sleduje v medicíně podrobněji určitou specifickou oblast, se již setkal s tím, že se stejným léčivým přípravkem bylo provedeno několik nezávislých, nicméně z hlediska designu, indikace a populace pacientů velmi podobných klinických studií, které však měly odlišný a někdy dokonce zcela protichůdný výsledek. Metaanalýzou klinických studií rozumíme metodiku sofistikovaného zobecnění výsledků dvou a více nezávisle provedených experimentů. Tato 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 v případě potvrzení homogenity jsou výsledky zobecněny. Cílem metaanalýz je přinést ucelenou a klinicky relevantní informaci na základě výsledků jednotlivých klinických studií, které mohou mít nejednotné závěry. Protože technologie metaanalýz je relativně složitá, není v tomto textu prostor zabývat se jí detailněji a uvedme si pouze několik málo příkladů, kdy může dojít k nežádoucímu zkreslení výsledků metaanalýz. Jedním ze základních problémů, se kterými se při metaanalýzách musíme vyrovnat je problém tzv. „publication bias“. Jedná se o to, že pokud pro metaanalýzu použijeme pouze výsledky klinických studií publikovaných v odborném tisku, můžeme snadno dojít k z hlediska pravdivé skutečnosti zkresleným závěrům, protože nezahrneme provedené, ale nepublikované studie. Je prokázáno, že studie, kterých se zúčastnil relativně nízký počet subjektů hodnocení, a které měly negativní výsledek (např. neprokázaly jednoznačný benefit nové léčby), mají menší šanci být publikovány v odborném tisku a to i přesto, že pokud byly provedeny metodicky správně, reprezentují pravdivou informaci.

Dalším, praktickým problémem při provádění metaanalýz, je problém vícenásobné publikace výsledků z experimentu provedeného na identické populaci subjektů hodnocení, což je důsledek faktu, kdy autoři po provedení náročného studie mají tendenci připravit z výsledků co nejvíce publikací často nazvaných a koncipovaných tak, aby nebylo jednoduché zjistit, že se jedná o tentýž projekt a tutéž populaci pacientů.

Zřejmě nejvýraznějším praktickým problémem metaanalýz je nejednotnost hodnocení účinnosti a bezpečnosti léčby v jednotlivých studiích, kdy např. přežití (celkové, doba do progresu apod.) může být v jedné studii měřeno od data diagnózy, v jiné od data randomizace nebo zahájení léčby apod., což představuje naprosto odlišné a obtížně srovnatelné údaje.

VELIKOST VZORKU (POČET SUBJEKTŮ HODNOCENÍ)

Počet subjektů hodnocení zařazených do řádně designovaného klinického hodnocení není náhodný a je přesně stanoven statistikem na základě tzv. „power analýzy“. Pokud je vzorek podhodnocen (do studie je tedy zařazeno méně subjektů než je třeba), je vysoce pravděpodobné, že daná studie nebude mít dostatečnou „sílu“ prokázat např. vyšší účinnost nebo bezpečnost nově hodnoceného přípravku i přes to, že tento přípravek skutečně účinnější a bezpečnější je. Naopak, pokud je do studie zařazeno více subjektů hodnocení než bylo potřeba, může se stát, že studie prokáže statisticky významně vyšší účinnost nebo bezpečnost nového přípravku, tento benefit však bude z hlediska klinického napros-

to bezvýznamný. Pro pochopení významu optimalizace velikosti vzorku v klinickém hodnocení je nutné správně interpretovat pojmy „klinicky významný rozdíl“ a „statisticky významný rozdíl“, které se někdy zaměňují ale ve skutečnosti znamenají diametrálně odlišné skutečnosti.

Klinicky významný rozdíl můžeme jednoduše definovat jako rozdíl v účinnosti a bezpečnosti nového přípravku oproti dosud standardně používanému, který je pro lékaře po zvážení všech ostatních okolností (např. bezpečnost, cena, náročnost aplikace apod.) dostatečný pro to, aby tento nový přípravek začal používat namísto dosud standardně používaného (pro zjednodušení zde neuvažujeme možnost individuálního přístupu k léčbě a možné rozdíly v benefitu léčby standardním a „novým“ přípravkem u různých pacientů). Klinicky významný rozdíl je tedy např. rozdíl v účinnosti (např. v podílu dosažení kompletních remisí nebo v době celkového přežití apod.) který definuje lékař na základě své expertní zkušenosti, bez ohledu na názor statistika. Klinicky významný rozdíl je tedy zcela nezávislý na provedení vlastního experimentu.

Naproti tomu *statisticky významný rozdíl* v účinnosti nebo bezpečnosti může být definován pouze na základě provedeního experimentu, a pokud tvrdíme, že např. nová léčba je statisticky významně účinnější, znamená to, že daný experiment prokázal, že existuje pouze velmi malá pravděpodobnost, že tento nový preparát je ve skutečnosti stejně nebo méně účinný než kontrola, avšak bez ohledu na to, „o kolik“ tento nový přípravek účinnější je. Obecně platí, že pro průkaz většího rozdílu v účinnosti potřebujeme v experimentu méně subjektů (menší vzorek) než pro průkaz menšího rozdílu. Při zařazení extrémně vysokého počtu subjektů do teoretické studie srovnávající celkové přežití pacientů užívajících odlišný typ udržovací léčby se nám tedy může stát (příklady byly publikovány), že prokážeme statisticky významný rozdíl v účinnosti léčby, který bude představovat rozdíl v přežití pouze jeden den, tento rozdíl v účinnosti bude zcela jistě považován za klinicky nevýznamný.

Cílem optimalizace velikosti vzorku v klinickém hodnocení je zařadit do projektu takový počet subjektů hodnocení, který bude dostatečný k tomu, aby při zjištění klinicky významného rozdílu v účinnosti a/nebo bezpečnosti mezi rameny studie byl tento rozdíl prokázán rovněž jako statisticky významný, tedy aby bylo možné říci, že pravděpodobnost, že tento výsledek byl s ohledem na počet zařazených subjektů hodnocení dosažen pouze náhodou, je minimální.

Důvody, proč je nutné optimalizovat velikost vzorku v klinických studiích je možné shrnout do následujících: *etické aspekty* (je neetické vystavovat působení experimentálního přípravku více subjektů než je nezbytně nutné), *ekonomické aspekty* (každý další subjekt zařazený do klinického hodnocení zvyšuje jeho náklady) a *požadavek legislativních autorit* (před zahájením klinického hodnocení musí být regulačním autoritám doložena metodika výpočtu velikosti vzorku, která je součástí protokolu každé studie).

ITT A PP POPULACE

V publikovaných výsledcích klinických studií se často setkáváme s konstatováním, že prezentované výsledky vznikly analýzou dat

Intention to treat nebo *Per protocol* populace pacientů. Pro vysvětlení těchto termínů použijme příkladu jednoduché hypotetické studie, ve které budeme srovnávat účinnost dvou léčebných postupů, kdy jeden z nich bude představovat např. radikální chirurgický výkon zatěžující významně pacienta a druhá léčebná modalita bude pouze medikamentózní léčba s minimální zátěží pacientů. Představme si, že do studie bude zařazeno celkem 200 pacientů, kteří podepíší informovaný souhlas a budou rovnoměrně v poměru 1:1 zařazeni do obou ramen studie. Kriteřiem pro zhodnocení účinnosti léčby bude v obou ramenech podíl pacientů s dosaženým určitým stupněm léčebné odpovědi. V rameni s medikamentózní léčbou bude tedy celkem 100 pacientů, z nichž všichni dodrží předepsané dávkování a podstoupí všechny úkony dané protokolem studie, při závěrečném zhodnocení bude u 50 (50 %) zjištěno dosažení požadovaného stupně léčebné odpovědi. V rameni s chirurgickou léčbou však celkem 20 (20 %) pacientů poté, co se dozvědí, že byli randomizováni do této skupiny odmítne podstoupit operaci (např. z obavy z rizik výkonu jako takového), operovaných bude tedy pouze 80 (80 %) pacientů a z nich u 40 bude konstatováno dosažení požadovaného stupně léčebné odpovědi. Pro srovnání účinnosti léčby v obou ramenech studie je tedy zásadní, zda budeme úspěšnost léčby v rameni s chirurgickým zákrokem hodnotit jako 40 úspěchů z 80 operovaných (úspěšnost 50 % a tedy srovnatelná s druhým ramenem) a nebo jako 40 úspěchů ze 100 plánovaných operací (tedy 40 % a tedy nižší účinnost než v kontrolním rameni). Řešení tohoto problému není jednoduché, protože oba přístupy nám podávají rozdílnou a vždy relevantní informaci. Jeden přístup svědčí o úspěšnosti dané metody jako takové a druhý o skutečné účinnosti metody při jejím zavedení do klinické praxe (pokud 20 % pacientů odmítne podstoupit operaci v průběhu studie, stejný, nebo spíše vyšší podíl pacientů odmítne podstoupit tento výkon i v průběhu klinické praxe). V případě, že pro hodnocení účinnosti a bezpečnosti léčby použijeme pouze soubor pacientů, kteří dodrželi léčebný protokol (tedy např. podstoupili daný chirurgický výkon), hovoříme o *Per protocol* analýze, přístupu, nebo populaci pacientů. V případě, že analýzu provedeme na souboru pacientů plánovaných pro použití daného léčebného schématu bez ohledu na to, zda bylo aplikováno nebo ne a v jaké míře, hovoříme o *Intention to treat* analýze, přístupu, nebo populaci pacientů. Při analýze klinických studií je nutné uvádět vždy výsledky zjištěné pro *Per protocol* i *Intention to treat* populaci a v ideálním případě se tyto výsledky zásadně neliší. Ve výše uvedeném příkladu bylo v důsledku zjednodušení relativně snadné odlišit, který konkrétní subjekt by byl zařazen do *Intention to treat* nebo jen do *Per protocol* populace (podle toho, zda podstoupil nebo nepodstoupil operační výkon). V praxi může být toto rozhodování komplikovanější, jako příklad uveďme odchylné od plánované dávky chemoterapie např. z důvodu snížení dávky při projevech toxicity, kdy je nutné stanovit určitou toleranci pro celkovou dávku přípravku, kdy až po jejím překročení je subjekt považován vyloučen z *Per protocol* populace.

LITERATURA

1. Bahit, M., C., Cannon, C., P., Antman, E., M. et al.: Direct comparison of characteristics, treatment, and outcomes of patients enrolled versus patients not enrolled in clinical trial at centers participating in the TIMI 9 Trial and TIMI 9 registry. *Am Heart J*, 145(1):109-117; 2003.
2. Hurwitz, H., Fehrenbacher, L., Novotny, W., et al.: Bevacizumab plus Irinotecan, Fluorouracil, and Leucovorin for the Treatment of Metastatic Colorectal Cancer. *N Engl J Med* 2004; 350:357-367.
3. Meinert, C., L.: *Clinical Trials: Design, Conduct, and Analysis*. Oxford University press, New York 1986.

4. Petitti, D., B.: *Meta-Analysis, Decision Analysis, and Cost- Effectiveness analysis*. Oxford Press, New York 2000.
5. Pocock, S., J.: *Clinical trials: A practical Approach*. John Wiley and Sons, New York 2000.
6. Rosenberger, W., F., Lachin, J., M.: *Randomization in clinical trials: Theory and practice*. John Wiley and Sons, New York 2002.
7. Spilker, B.: *Guide to Clinical Trials*. Raven Press, New York 1991.
8. Twelves, Ch., Wong, A., Nowacki, M., et al.: Capecitabine as Adjuvant Treatment for Stage III Colon Cancer. *N Engl J Med* 2005; 352:2696-704.
9. Weijer, Ch.: Are clinical trial participants comparable to non-participants? *Control Clin Trials*. 16:48S; 1995.

ZÁKLADNÍ TYPY PARAMETRŮ HODNOCENÝCH V KLINICKÝCH STUDIÍCH V ONKOLOGII

PRIMARY END-POINTS IN CLINICAL TRIALS IN ONCOLOGY

DEMLOVÁ R.

ODDĚLENÍ KLINICKÝCH HODNOCENÍ, MASARYKŮV ONKOLOGICKÝ ÚSTAV, BRNO

Souhrn: Cílem tohoto sdělení je přiblížení a výklad jednotlivých parametrů hodnocení účinnosti v klinických studiích v onkologii, zvážení vhodnosti jejich použití v exploračních nebo konfirmačních studiích a zamýšlení se nad problémy plánování klinických studií v případě nové, tzv. cílené biologické léčby.

Klíčová slova: explorační studie, konfirmační studie, response rate, doba do progresu, přežití, kvalita života

Summary: This article will review commonly used clinical trial end-points and describe their potential advantages or disadvantages in exploratory or confirmatory clinical trials in oncology. This review also discusses the problems with study design in case of new targeted biological therapy.

Key words: exploratory trials, confirmatory trials, response rate, time to progression, survival, quality of life

Účinnost nového léčiva nebo nové léčebné kombinace včetně průkazu bezpečnosti musí být před jeho registrací a uvedením na trh prokázána v rozsáhlých kontrolovaných studiích, které jsou naplánovány a prováděny v souladu se všemi celosvětově uznávanými pravidly pro provádění klinických hodnocení. O těchto pravidlech i dalších náležitostech kolem klinických hodnocení pojednávají další články uvedené v tomto supplementu. Výsledkem takto naplánovaných studií by pak mělo být jednoznačné definování cílové skupiny pacientů, pro kterou má nová léčba největší přínos, zhodnocení bezpečnostního profilu včetně uvedení očekávaných nežádoucích účinků léčiva a vyhodnocení poměru přínosu a rizik při použití nového léčiva v širší klinické praxi.

Celý proces klinického hodnocení je tradičně dělen na IV fáze, přičemž fáze I - III předcházejí případně registraci léčiva. Z hlediska informačního lze studie méně tradičně rozdělit na studie popisné, explorační a studie konfirmační.

Explorační studie nemusejí vycházet z testované hypotézy, mohou být vedeny retrospektivně nebo prospektivně, mohou ale nemusí srovnávat dvě nebo více skupin. Informace získané z těchto exploračních fází klinického výzkumu pak mohou přispět a často i výrazně napomoci zhodnocení efektu i bezpečnosti nového léčiva. Toto se ukazuje velmi přínosné zejména v posledních letech, tedy v době zavádění nových cílených molekul neboli tzv. biologické léčby nádorů, kdy je se zvyšuje potřeba odhalit ty „správné“ prediktivní faktory a biologické markery pro účinek takto působícího léčiva. Faktum zůstává, že průkazné závěry lze pak učinit až na základě optimalizovaných randomizovaných konfirmačních studií, které jsou také uznávány všemi regulačními orgány při závěrečném posuzování nového léčiva.

Konfirmační studie představují kontrolovaný optimalizovaný systém založený na srovnání různých léčebných postupů, v onkologických studiích většinou současného standardního a experimentálního ramene léčby. Protokolem studie jsou pak dána veškerá pravidla provedení klinického hodnocení včetně uvedení primárního cíle studie (s vhodnými cílovými parametry), definováním skupiny vhodných pacientů (neboli sub-

jektů klinického hodnocení, použijeme-li správnou terminologii), definováním nulové hypotézy, statistické analýzy a léčebného plánu včetně sběru dat.

Následující text by měl být věnován především jednotlivým cílovým parametrům, které se objevují jak v exploračních, tak i konfirmačních klinických studiích [1,5,6].

K tomu, abychom byli schopni vyhodnotit účinnost nového léčiva nebo nové kombinace léčiv již v dřívější fázi klinického hodnocení a odhadnout tak jeho nebo jejich biologický efekt již v samém počátku klinického zkoušení, používáme při plánování protokolu klinické studie tzv. „surrogate endpoints“, říkejme jim jakési *zástupné nebo náhradní cílové parametry*. Zhodnocením těchto parametrů bychom měli být schopni predikovat klinický přínos nového léčiva, přičemž musíme mít stále na paměti, že závěrečné hodnocení celkového přínosu daného léčiva či nové léčebné kombinace musí vycházet z přísně naplánované konfirmační studie, nebo používá známější terminologii, ze srovnávací studie III. fáze.

Response Rate (RR)

Nejčastěji používaným parametrem ve smyslu „surrogate endpoint“ je tzv. Response Rate (RR). Response Rate je možno definovat jako procentuální dosažení kompletních nebo parciálních odpovědí na podávanou léčbu, a to v protokolu předem definovaných časových intervalech. Obecně lze říci, že Response Rate je považován za vhodný parametr při objektivizaci protinádorové aktivity, neboť spontánní regrese nádoru je výjimečná a zmenšení nebo vymizení nádoru pak lze jednoznačně přičíst léčebnému efektu podávané léčby.

Při hodnocení odpovědí jsou v klinických studiích nyní nejčastěji používána tzv. RECIST kriteria (Response Evaluation Criteria in Solid Tumors), o kterých bude pojednáno v samostatném sdělení. Stejně jak u WHO kriterii i RECIST kriteria rozlišují kompletní a parciální odpověď (Complete Response, Partial Response - CR, PR), stabilizaci onemocnění (Stable Disease -SD) a progresi (Progression - PD). Response Rate je tedy definován jako procento kompletních nebo parciálních odpovědí (CR + PR), stabilizace onemocnění není hodnoceno jako odpověď [2].

Nevýhodou hodnocení tohoto parametru mohou být nepřesnosti při použití různých radiodiagnostických metod nebo častěji spíše nepřesnosti při radiologickém čtení a měření odpovědi, zejména pokud se jedná o multicentrické studie. Stále častěji se tak v klinických studiích objevuje role tzv. centrálního radiologa a role nezávislé hodnotící komise (Independent Radiological Review Committee), jejichž závěry jsou pro vyhodnocení odpovědi určující. Nezbytností je pak zasilání všech radiologických nálezů, což samozřejmě navyšuje administrativu a také finanční náklady na studii.

Neméně důležitým parametrem je také doba trvání odpovědi, která přináší komplexnější pohled ve srovnání s izolovaným nahlížením na samotný Response Rate. Opět je nutno zmínit, že doba trvání odpovědi je závislá na četnosti hodnocení odpovědi nádoru, respektive na časových intervalech mezi jednotlivými hodnoceními (v klinických studiích nejčastěji 8-12 týdnů). Při srovnávání výsledků různých studií je na tuto skutečnost nutno pamatovat!

TTP (Time to Progression)

V konfirmačních studiích v onkologii je často jako primární srovnávaný parametr používán tzv. Time to Progression - doba do progresse onemocnění. Výhodou tohoto parametru je především to, že jej lze dosledovat v kratším časovém období než je tomu u parametru celkového přežití. Oproti celkovému přežití rovněž není jeho výsledek ovlivněn nebo zkreslen podáním druhé, případně další linie léčby. Faktem ale zůstává, že přesné určení doby do progresse onemocnění je vždy určitým odhadem, neboť k aktuální progresi dochází mezi dvěma plánovanými hodnoceními, nikoliv přesně v den provedení vyšetření, při kterém byla progresse zaznamenána. Z toho také vyplývá, že četnost jednotlivých vyšetření, respektive jejich časový odstup opět mohou tento sledovaný parametr ovlivnit. Je zcela zásadní, aby jednotlivá vyšetření byla v obou srovnávaných ramenech léčby plánována naprosto stejně a je třeba opět podotknout, že je nutno si tuto skutečnost uvědomit při srovnávání výsledků různých studií. Precizní hodnocení klade vysoké nároky na samotné provedení daného vyšetření (nejčastěji CT nebo MRI), zhodnocení nálezů a nutnost dodržení vždy stejné vyšetřovací metody po celou dobu sledování konkrétního pacienta. Z tohoto důvodu je také tento parametr, tedy TTP, označován za časově i finančně nejnáročnější parametr, zejména ve srovnání s parametrem přežití.

Problém v onkologických studiích může nastat také v tom, co vlastně považujeme nebo „klinicky cítíme“ jako progresi onemocnění. Protokolem studie podle RECIST kritérií popisovaných v jiné kapitole tohoto suplementa je to jednoznačně objevení se nové léze nebo minimálně 20% progresse měřitelných lézí. V klinické praxi se však můžeme setkat s dalšími „příznaky“ pravděpodobné progresse onemocnění, ať již například posunem v nádorových markerech, objevením se nové metabolické aktivity při stále dostupnějším PET vyšetření nebo i klinických známkách jako zhoršení celkového Performance Status či objevení se nebo zhoršení symptomů souvisejících s primárním nádorem nebo jeho metastázami. Toto jsou klinické situace, se kterými většinou protokol studie nepracuje, nicméně není možno s nimi nepočítat.

Pokud se při následné statistické analýze prokáže statisticky signifikantní prodloužení doby do progresse onemocnění v jednom ze srovnávaných ramen, bereme tento výsledek jako průkazný. Trošku složitější je to pak s převedením „statistické významnosti“ do „významnosti klinické“. Je otázkou, zda například 8-týdenní prodloužení doby do progresse onemocnění přináší jasný klinický benefit pro asymptomatického pacienta léčeného s paliativním záměrem, nemluvě o finanční nákladnosti nových léčiv.

OS (Overall Survival), DFS (Disease-Free Survival)

Prodloužení přežití u pacientů s inoperabilním lokálně pokročilým nebo metastazujícím nádorem je v konfirmačních stu-

diích, srovnávajících dvě ramena léčby, považováno za „nejprokazatelnější“ cílový parametr. Srovnatelným parametrem v případě adjuvantní léčby je pak prodloužení přežití bez známek onemocnění (disease-free survival).

Oba patří mezi parametry, které se dají vyhodnotit zcela přesně, a to i bez ohledu na jednotlivá měření nádorových ložisek. Současně je možno říci, že patří mezi parametry, které jsou schopny jednoznačně vyhodnotit případný poměr rizika a benefitu při zavádění nového léčiva do klinické praxe. Na druhou stranu však není možno nezmínit i nevýhody použití tohoto parametru jako primárního hodnoceného cíle. Nejzávažnějším negativním argumentem je především ovlivnění celkového přežití následnými léčebnými postupy. Příkladem, je-li nové léčivo srovnáváno se standardní léčbou v první linii metastazujícího nádoru kolorekta, pak je celkové přežití v obou srovnávaných ramen ovlivněno druhou, případně dalšími liniemi léčby. Protokoly klinických studií se snaží řešit tento problém tzv. zkřížením (cross-over) obou ramen léčby po zaznamenané progresi, což následně umožňuje vyhodnotit skutečný přínos nového léčiva i v ovlivnění celkového přežití.

Přežití je parametr, jehož sledování je závislé na čase. Obecně jsou analýzy tohoto typu nazývány „time-to-event analyses“, neboť po určitou dobu sledování (tzv. follow-up time) zaznamenáváme postupně rostoucí počet hodnocených událostí. Definiční významných parametrů hodnocených analýzou přežití v onkologii přináší tabulka 1. V závislosti na typu onkologické diagnózy a klinickém stádiu onemocnění může ovšem doba sledování nutná pro zachycení většího počtu událostí představovat měsíce až roky. Dostáváme se tak do problému s průběžným hodnocením přežití, neboť rozhodneme-li se hodnotit přežití v nějakém konkrétním bodě, nebude čas nastání hodnocené události znám pro všechny zařazené subjekty. V konkrétním čase trvání studie získáváme jednak subjekty, u kterých již událost proběhla a její čas je znám, a subjekty bez události, dále pokračující ve studii - doba sledování těchto případů je k datu analýzy brána jako cenzorovaná („censored times“).

Přežití lze srovnávat například u dvou nebo více experimentálních skupin pacientů v určité studii nebo i v rámci jednoho souboru pacientů tříděného na podskupiny. V druhém případě hovoříme o tzv. stratifikaci souboru, stratifikované analýze přežití a parametr dělicí soubor pacientů na podskupiny je nazýván stratifikační faktor. Mezi významné stratifikační faktory v onkologii patří klinické stádium, TNM klasifikace, věk nebo výsledky předchozí léčby.

Podrobněji se lze o analýze přežití informovat například v práci Altman, 1991 [3].

Quality of Life (QoL)

Častým parametrem hodnoceným v konfirmačních studiích bývá také zhodnocení kvality života. Prospektivně získané informace o kvalitě života získané přímo od pacienta jsou velmi cenné, neboť přináší jeho vlastní pohled na případný klinický benefit podávané léčby, ať již hodnocením symptomů onemocnění, případných nežádoucích účinků spojených s léčbou nebo ovlivněním běžných denních aktivit. Hodnocení kvality života je samozřejmě velmi komplikovanou záležitostí, neboť je ovlivněno subjektivními pohledy jednotlivých pacientů. Současně bývá často ovlivněno stupněm compliance pacienta, není výjimkou, že řada informací v dotaznících zůstává nevyplněna. Kvalita života jako hodnocený parametr by měla být použita pouze u prospektivně plánovaných, randomizovaných a zaslepených studiích tak, aby se minimalizovalo případné bias.

Všechny předchozí popsané hodnocené parametry, používané v exploračních nebo konfirmačních studiích, lze zjednodušeně označit jako parametry tradiční, používané v onkologických studiích již dlouhou dobu. V posledních několika letech se do klinického zkoušení dostala řada nových léčiv, tzv. „tar-

geted“ molekuly, působících cíleně na struktury na buněčné nebo molekulární úrovni. Může se jednat příkladem o malé molekuly ovlivňující kinázové podjednotky růstových receptorů, monoklonální protilátky nebo antisense oligopeptidy. Nejedná se tedy o klasicky cytotoxicky působící léčiva tak, jak jsme byli zvyklí u dlouhodobě zaužívaných cytostatik. Všechna tato nová léčiva vyžadují již z principu jejich mechanismu účinku i pozměněný přístup k definování těch nejvhodnějších cílových parametrů, které by v klinických studiích s těmito léčivy vedly k tomu nejspříhodnějšímu zhodnocení jejich efektu včetně odhalení určitých prediktivních markerů, které by nás vedly k použití tohoto nového léčiva u „správneho pacienta ve správnou dobu“. O léčbě „šité na míru“ už toho bylo hodně řečeno, hodně popsáno, nicméně řada otázek zůstává stále nedořešena. Jednou z nich je také potřeba nového přístupu v designování klinických studií u tohoto typu léčby [4].

Kdybychom postupovali v tradičním pojetí rozdělení studií do jednotlivých fází, pak by pravděpodobně v první fázi klinického zkoušení mělo být cílem definovat dávku nebo koncentraci léčiva, která vede k maximálně účinné inhibici cílového receptoru, nikoliv dosažení maximálně tolerované dávky na základě hodnocení toxicity. Ve fázi II bychom měli být schopni správně vyhodnotit účinnost nového léčiva, přičemž použití klasického endpointu response rate se nejeví jako nejvhodnější už proto, že řada těchto léčiv nepůsobí zmenšením masy nádorů. Nové léčivo by mělo být schopno ovlivnit a změnit biologické chování vlastního nádoru a v klinické studii bychom pak měli hodnotit zejména vhodné biomarkery, navíc pouze v populaci pacientů, u kterých se jeví použití nového léčiva jako vhodné (například na základě stanovení exprese daného receptoru). Designování konfirmačních studií III. fáze tak, abychom potvrdili jednoznačný klinický přínos nového léčiva, zůstává stále velkou výzvou.

Tabulka č. 1 */

Parametr	Definice
Celkové přežití (Overall survival, OS)	Časový interval mezi randomizací pacienta a smrtí. Žijící pacienti jsou cenzorováni k datu poslední kontrolní návštěvy. Analýza může být dle okolností omezena pouze na pacienty s určitou příčinou smrti (primární onemocnění, jiné).
Bezpříznakové přežití (disease-free survival, relaps-free survival, DFS)	Časový interval mezi randomizací pacienta a prvním projevem další aktivity nemoci (relaps) nebo úmrtím. Hodnocení se týká pouze pacientů, kteří dosáhli komplexní remise. Pacienti v kompletní remisi, kteří zemřou před projevem aktivity choroby, jsou plně hodnoceni k datu úmrtí. Smrt pacienta ještě před relapsem je v tomto parametru brána jako projevená událost („event“) a ovlivňuje profil přežití. Žijící pacienti bez příznaků nemoci jsou cenzorováni k datu poslední kontrolní návštěvy.
Doba do progresse (progression-free interval, time to progression, TTP)	Časový interval mezi randomizací pacienta a progresí onemocnění. Pacienti, kteří zemřou před projevem progresse jsou cenzorováni k datu úmrtí. Žijící pacienti bez progresse onemocnění jsou cenzorováni k datu poslední kontrolní návštěvy. Při hodnocení TTP nemusí pacient žít zcela bez příznaků choroby, na rozdíl od DFS může být například ve stavu stabilizovaného onemocnění.
Interval přežití bez příznaků nemoci (Disease free interval, DFI)	Časový interval mezi randomizací pacienta a relapsem. Týká se pouze pacientů v kompletní remisi. Pacienti, kteří zemřou před projevem relapsu jsou cenzorováni k datu úmrtí. Žijící pacienti bez příznaků nemoci jsou cenzorováni k datu poslední kontrolní návštěvy.
Jakýkoli „time-to-event“ interval (event-free survival, EFS)	Protokolárně specifické parametry přežití, které definují různé události („events“) ve vývoji pacienta. Při definici je třeba dbát na způsob zařazení pacientů zemřelých bez projevu této události, při zařazení půjde o hodnocení přežití, při cenzorování těchto pacientů půjde o hodnocení časového intervalu k dané události.
Sumarizované časové intervaly	
Doba trvání kompletní remise	Časový interval mezi randomizací pacienta a projevem aktivity nemoci (relaps). Tato charakteristika je vyhodnocována pouze pro pacienty, kteří po určité léčebné fázi dosáhli kompletní remise. V případě celkového hodnocení pacienta, který prošel více fázemi léčby, například při relapsech onemocnění, je možné hodnotit celkovou sumární dobu setrvání v kompletní remisi.
Doba trvání odpovědi	Stejný parametr jako doba trvání kompletní remise, je pouze definičně rozšířen na pacienty s kompletní i částečnou remisi onemocnění.
Výše uvedené definice obsahují vždy jako vstupní datum pro hodnocení pacienta randomizaci v rámci klinické studie. Pro běžnou klinickou praxi bez konfirmačních protokolů toto datum nahrazuje datum zahájení primární léčby, případně účelově jiné datum závažného zlomu ve vývoji léčby pacienta (např. datum transplantace kostní dřeně, apod.). Blíže viz též [8].	

*/ převzato : L. Dušek, Analýza klinických dat v onkologii

Literatura :

- Schilsky R.L. End Points in Cancer Clinical Trials and the Drug approval Process, Clin Cancer Res Vol. 8:935-938, 2002
- EORTC (European Organization for Research and Treatment of Cancer) (1996). A Practical Guide to EORTC Studies. Therass E P. (editor in chief), 227p., EORTC Data Center, Brussels, Belgium. ISBN 2-930064-06-4.
- Altman D.G. (1991) Practical Statistics for Medical Research. 611p., Chapman and Hall, London.

- Korn E.L., Arbuck S.G., Pluda J.M., Simon R., Kaplan R.S., Christian M.C. Clinical Trial Design for cytostatic agents - are new approaches needed? J Clin Oncol, 19:265-272, 2001
- Pazdur R. Response Rate, Survival and Chemotherapy Trials. Journal of the NCI, 19: 1552-1553, 2000.
- Holland J.F. Clinical Trials in Cancer. Clinical Cancer Research, 3: 2585-2586, 1997.

DESIGN KLINICKÉ STUDIE

DESIGN OF A CLINICAL TRIAL

TOMÁŠEK J., KISS I.

ODDĚLENÍ KLINICKÉ ONKOLOGIE FN BRNO

Souhrn: Randomizované klinické studie poskytují nejpřesvědčivější důkazy o účinnosti léčby. Observační studie jsou méně průkazné proti experimentálním. Je zde velké riziko zkreslení, pokud není jasný vztah mezi působícím faktorem a výsledkem.

Klíčová slova: design klinické studie

Summary: The study design that provides the most compelling evidence of the treatment effect is the randomized controlled trial. Observational studies provide weaker empirical evidence than do experimental studies because of the potential for large biases to be present when there is an unknown association between a factor and an outcome.

Key words : clinical trial design

Vědecké poznání je snaha o zjištění pravdivé skutečnosti v určité oblasti výzkumu. Novověká věda je založena na představě zákonitého chování skutečnosti. V oblasti přírodních věd je hypotéza, doposud neprokázané tvrzení, ověřována pozorováním a opakovatelným experimentem. Při pozorování je studován děj nijak neovlivněný výzkumníkem. V oblasti medicíny jde například o problematiku diagnostiky, kdy je hodnocen vztah příznak - nemoc. Experimentální metoda je aktivnější, v medicíně zahrnuje hlavně sledování efektu terapeutických intervencí. Patří sem ale i například experimentální studium fyziologických a patofyziologických dějů. Dobře vedená klinická studie je nejbezpečnější a nejrychlejší způsob hledání nových léčebných možností. Design (uspořádání) klinické studie určuje plán, jakým bude studie prováděna a jak tedy budou získávána data pro statistickou analýzu určitého jevu. Design studie je velmi důležitý a z určitého pohledu významnější než následná statistická analýza. Špatný design studie nemůže být zpětně opraven, ale špatně analyzovaná data studie mohou být vždy podrobena analýze nové. Klinické studie můžeme rozdělit podle designu na 2 hlavní skupiny: studie observační a studie intervenční, z nichž nejvyšší informační hodnotu mají studie experimentální (konfirmační).

Observační (popisné) klinické studie:

V observační studii není přičleňování sledovaných faktorů subjektům studie nijak ovlivněno výzkumníkem, jde v podstatě o „přirozený experiment“. V případě zkoumání vlivu některých faktorů, jejichž záměrné přičleňování subjektům hodnocení by bylo neetické, zbývá pouze tento typ studie. Observační studie obecně poskytují slabší empirické důkazy než studie experimentální. Je zde větší možnost zmatečných zkreslení skutečnosti v případě, že není jasný vztah mezi sledovaným faktorem a výsledkem. Význam těchto studií spočívá také v předběžných výsledcích, které mohou být podkladem designování přesné experimentální studie.

Kohortní studie („cohort study“). Prospektivně je sledován soubor (kohorta) osob, z nichž část byla, je nebo bude vystavena působení určitého faktoru. Po uplynutí doby potřebné ke vzniku nějakého významného jevu (např. vznik nemoci, kom-

plikace nemoci, relaps onemocnění, úmrtí) je zhodnocen výskyt tohoto jevu ve skupině ovlivněné studovaným faktorem a ve skupině bez něj. V kohortní studii se postupuje od uvažované příčiny k následku, je tak možno určit kauzální vztah mezi určitým faktorem a následným jevem.

Výběr subjektů do kohorty a parametry, které tento soubor definují, podstatně ovlivňují informační hodnotu výsledku studie.

Studie „případ - kontrola“ („case-control study“). Při tomto plánu je postup obrácený proti prospektivnímu sledování. Retrospektivně je sledován soubor osob u nichž došlo ke sledovanému jevu (např. vznik onemocnění) a hodnotíme výskyt určitého faktoru (většinou rizikového nebo ochranného) v minulosti. Údaje jsou hledány většinou ve zdravotní dokumentaci. Kontrolní skupinu tvoří retrospektivně dohledané osoby bez sledovaného jevu.

Tento typ klinické studie je často využíván ke studiu rizikového nebo ochranného vlivu nějakého faktoru (např. význam pasivního kouření pro vznik nádorového onemocnění, protektivní účinek ovoce a zeleniny nebo vliv dlouhodobého užívání některých léků na vznik onemocnění). Dalším případem jsou vzácné diagnózy a nemoci, jejichž vývoj trvá mnoho let a prospektivní sledování by bylo těžko uskutečnitelné. Informační hodnota studie je dána výběrem případové a kontrolní skupiny. Problematická je možná absence potřebných údajů ve zdrojové dokumentaci, je riziko vstupu chybných údajů. Obtížné je vytvoření kvalitní kontrolní skupiny. Je třeba vždy řešit problém, zda má jít o jedince zdravé nebo jedince s příbuzným onemocněním a pod.

Průřezová studie (prevalenční, cross-sectional study). Jde o observační studii popisující vztah mezi určitým jevem (nemocí) a některými faktory v definované populaci obvykle v jednom časovém okamžiku. Tyto studie nezkoumají vůbec vztah mezi délkou působení faktoru a touto nemocí. Tyto studie jsou relativně levné a mají význam, pokud se sledovaný faktor nemění v čase (např. pohlaví, krevní skupina). V průřezové studii je často těžké rozlišit mezi příčinou a následkem. Statisticky významným výstupem studie je zjištění prevalence sledovaného jevu.

Série kazuistik (case series). Popisná studie série případů obvykle popisuje projevy, klinický průběh a prognózu určitého stavu. Informační význam takové série je velmi malý, pokud se výsledky sledování podstatně liší od předpokládaného výsledku. Chybí zde jakékoliv porovnání s kontrolním souborem. Tento typ souborů je často kliniky prezentován v odborné literatuře, informační hodnota výsledků může být velmi sporná, údaje je nutno vnímat velmi kriticky. Výsledky takového sledování mohou být zdrojem hypotéz pro návrh věrohodněji uspořádané studie.

Kazuistika (case report). Jde o prezentaci jednoho případu obvykle opět popisující projevy, klinický průběh a výsledek léčby. Díky obrovské přirozené biologické variabilitě je důkazní hodnota kazuistiky minimální.

V epidemiologii bývá často volena **ekologická studie (ecologic study, aggregate study)**. Ve studii tohoto typu není charakteristika jednotlivce vůbec získána, v rámci skupin osob je vyhodnocován vztah mezi nemocí a předpokládaným rizikovým faktorem. Jednotkou analýzy není jedinec, ale skupina jedinců. Nejčastějším důvodem volby této studie je situace, kdy data na individuální úrovni nejsou k dispozici. V ekologických studiích se výsledky vztahují ke sledovaným skupinám a nikoliv k jednotlivcům. Závěry studie nejsou také platné ve vztahu k jednotlivci. Odlišnost asociací na skupinové a individuální úrovni bývá označována jako ekologické zkreslení či ekologický klam (ecological fallacy) a představuje hlavní omezení ekologických studií. Členům skupiny jsou přiřazovány charakteristiky, které však jednotlivci nevykazují. Příkladem takové studie je sledování mortality na karcinom plic v závislosti na spotřebě cigaret na hlavu ve zkoumané populaci.

Intervenční studie

Intervenční studie bez kontrolního ramene jsou častým designem pilotních projektů, kdy je testován spíše mechanismus účinku léku, jeho farmakokinetika a farmakodynamika. Jsou používány ve studiích fáze I a často i II. Používají se také v případech, že neexistuje srovnatelný postup alternativní, nebo by byl jiný postup neetický.

Intervenční nerandomizované studie. Nedochozí zde k náhodnému přiřazování (alokaci) pacientů do jednotlivých léčebných ramen. Léčba je volena na základě přání pacienta nebo rozhodnutí lékaře. Je zřejmé, že parametry subjektů v různých ramenech se mohou podstatně odlišovat, není rovnoměrné rozložení rizikových faktorů v ramenech a dochází tak ke zkreslení výsledků. Nerandomizovaná studie je zcela nevhodná ke srovnávání různých léčebných postupů.

Experimentální (konfirmační) studie:

Experimentální studie jsou oproti popisným studiím plně kontrolovány, jsou zde obvykle porovnávány 2 nebo více léčebných metod na skupinách subjektů - pacientů. Rozdělování pacientů do jednotlivých léčebných skupin probíhá pod kontrolou, je tedy možná randomizace. Při vybalancovaném uspořádání studie randomizace probíhá tak, aby byla plně srovnatelná všechna ramena studie symetricky ve všech podstatných faktorech, které by mohly i jen teoreticky ovlivnit průběh a výsledek studie. Dobře designovaná a řádně provedená experimentální klinická studie poskytuje nejvyšší míru objektivních a reprodukovatelných závěrů ohledně bezpečnosti zkoumaného léčebného postupu a jeho efektivity. Možnost zkreslení a náhodných chyb je tímto snížena na minimum.

Základem každé experimentální klinické studie je dobře formulovaná hypotéza. V hypotéze se řeší předpokládaný vztah výsledku experimentálního ramene proti kontrole. Jde o prokázání, že léčba je lepší než v kontrolním rameni (superiorita), že není horší než v kontrolním rameni (non-inferiorita), že výsledky jsou stejné (ekvivalence). V hypotéze lze také hle-

dat závislost účinku na dávce léku. V každé studii musí být stanoveny cíle (endpoints). Primární cíl rozhoduje o výsledku studie. V onkologii bývá nejčastějším primárním cílem odpověď na léčbu a hodnocení přežití. Sekundárním cílem pak bývá hodnocení toxicity léčby, nežádoucí účinky, kvalita života nebo ekonomické aspekty léčby.

Cíle studií mohou procházet na základě analýz již publikovaných studií vývojem. Příkladem mohou být studie zaměřené na adjuvantní chemoterapii kolorektálního karcinomu.

Primárním cílem takto zaměřených studií bývá hodnocení podílu pacientů, kteří přežijí 5 let - pětileté přežití. Tyto studie jsou tedy časově velmi náročné a trvá dlouhou dobu než se výsledky dostanou do praxe. Zhodnocením údajů 12 915 pacientů z 15 velkých randomizovaných studií III. fáze s režimy na bázi fluorouracilu bylo dokázáno, že tříleté přežití bez známek onemocnění s vysokou pravděpodobností predikuje pětileté přežití (Sargent DJ 2004). Je tak možné u obdobných studií stanovit jako primární cíl právě tříleté přežití bez známky onemocnění a konečné výsledky získat o 2 roky dříve.

Podle formulované hypotézy a cílů studie je nutné volit optimální počet pacientů zahrnutých do projektu a je vybrán vhodný způsob statistické analýzy. Hodnotu závěrů klinické studie významným způsobem ovlivňuje reprezentativnost studované skupiny jedinců. Toto ovlivňují vstupní a vylučovací kritéria. U studií s více rameny pak proces randomizace zajistí rozdělení subjektů do skupin tak, aby byly tyto skupiny srovnatelné i v parametrech, které nejsou přímo sledovány, ale které mohou ovlivnit výsledek hodnocení. Konkrétní randomizační plán je popsán v protokolu studie. Je zřejmé, že na návrhu plánu studie se musí podílet statistik.

Randomizované klinické studie lze rozdělit podle designu do několika hlavních typů.

Při **paralelním designu** jsou srovnávána dvě ramena studie, případně je těchto ramen více.

Při **zkříženém designu** (cross-over) jsou pacienti také randomizováni do 2 ramen. V první části studie je podána zkoumaná léčba pacientům v jednom rameni a pacientům v druhém rameni je podáno placebo nebo jiná referenční terapie. V následné druhé části se pak léčba mezi těmito rameny vymění. Každý pacient zařazený do studie je tedy léčen oběma léčebnými postupy, o pořadí těchto postupů rozhoduje vstupní randomizace. Porovnává se tedy účinek 2 léčebných postupů na stejných pacientech. Mezi oba léčebné postupy je třeba vložit určitý časový interval (wash out period), aby přestal působit lék podávaný v prvním období studie.

Kladnou stránkou tohoto uspořádání studie je porovnávání účinku léčby na stejných pacientech, při paralelním uspořádání je porovnáván účinek na různých osobách. K prokázání rozdílných výsledků vyžadují studie se zkříženým designem obecně menší počet osob než studie ostatní. Mezi hlavní úskalí takovéto studie patří možnost přenesení nežádoucích účinků z první části studie do části druhé (carry-over effect), ovlivnění výsledku léčby pořadím podávaných léků (order effect) a možnost změny charakteristiky pacientů v průběhu studie.

Faktoriální intervenční studie umožňuje zkoumání účinku několika léků současně. Musí být přitom možné podávat tyto léky současně, nesmí spolu nijak interferovat, účinek jednoho léku nesmí být závislý od druhého, nesmí docházet při současném podání ke zvýšené toxicitě. Při tomto uspořádání studie jsou jednotlivým skupinám sledovaných pacientů podávány všechny možné kombinace dvou (či více) zkoumaných léčebných postupů. Příkladem je zkoumání účinnosti léku A proti kontrole bez léčby a léku B proti kontrole bez léčby. Randomizací bude mít 25% pacientů ve studii lék A, 25% lék B, 25% bude léčeno kombinací AB a 25% pacientů bude bez léčby. Faktoriální uspořádání umožňuje testovat současně účinek 2 a více léků bez zvýšených nároků na počet zařazených pacientů. Toto uspořádání studie je velmi flexibilní. Výsledek testů

vání jednoho léku může být například znám dříve a podle něj lze tento lék všem pacientům ve studii vysadit nebo nově podat. Nijak tím není ovlivněn průběh testování léku B.

Zaslepení studie a typy kontrol

Výsledky intervenčních studií mohou být ovlivněny subjektivními faktory jak ze strany pacienta (např. přehnané očekávání nebo strach z nového léku), tak ze strany zkoušejícího (např. očekávání lepších výsledků v některém rameni, strach z vedlejších účinků léku).

Zaslepením studie lze možnost tohoto zkreslení minimalizovat. Pokud pacient neví, jakým způsobem je v rámci studie léčen, jde o **jednoduchou slepou studii**. Ve **dvoujité slepé studii** tuto informaci nemá ani zkoušející, v **trojité slepé studii** pak ani osoba provádějící statistické zhodnocení studie. Zvláštním typem studie je metoda „**double dummy**“, česky nejspíše „dvojitý zástup“. Potřebujeme-li porovnat účinnost dvou různých lékových forem jednoho léku, například antiemetikum podávané perorálně v tabletách se stejným antiemetikem podávaným intravenózně, jsou pacientům v obou ramenech studie tohoto designu podávány vždy infuze a tablety společně. V jednom z ramen je aktivní látka jen v perorální formě a v druhém rameni pouze v intravenózním přípravku.

V některých případech klinického hodnocení není zaslepení studie možné, hovoříme pak o **otevřené studii**. Je tomu tak například při velkých rozdílech ve vedlejších účincích léků v aktivním a v kontrolním rameni. Někdy nejde vyrobit srovnatelné placebo (např. chuťově typický nutriční přípravek). V otevřených studiích jsou také ověřovány nové chirurgické postupy.

Při návrhu klinické studie je vždy nutné dobře volit léčbu v kontrolním rameni. V placebem kontrolované studii ověřujeme, že účinek zkoumané látky je od placeba rozdílný a že látka je

skutečně specificky účinná. Srovnáváme-li ji s referenční léčbou („komparátorem“), obvykle očekáváme, že najdeme jen nepatrný nebo žádný rozdíl oproti již dostupné, obvykle nejlepší standardní léčbě. Obě ramena také mohou být aktivní. Je porovnáván např. účinek dvou chemoterapeutických kombinací, jejichž účinek byl již ověřen jinými studiemi.

Příkladem **externí kontroly** je **historická kontrolní skupina**. Data o této skupině jsou získána ze zdravotní dokumentace, případně jsou používány údaje literární. Výhodou je, že žádnému pacientovi není odeprána „nová“ léčba. Je to také jediný možný postup v případě, že srovnávacím ramenem je léčba již zastaralá, v současné době non lege artis (např. pouze podpůrná terapie nádorů, kde již existuje účinná protinádorová léčba). Nábor do takto koncipované studie je jednodušší a studie je levnější ve srovnání s randomizovanými studiemi.

Použití historické kontroly je však zatíženo rizikem velkého zkreslení. Jde například o možné změny v diagnostice zkoumané nemoci, výsledky ovlivní také jiná podpůrná terapie. Pokud je používána jako kontrolní rameno skupina z jiného zdravotního zařízení nebo dokonce jiného státu, je zde možnost dalšího zkreslení. Výsledky jsou ovlivněny obecně odlišnými diagnostickými a terapeutickými postupy, odlišnou úrovní ve vybavení zdravotnického zařízení a také rozdíly v populacích pacientů.

Při plánování a hodnocení výsledků klinické studie je konzultace biostatistika nezbytná. V případě velkých randomizovaných studií je biostatistik většinou přímo v řešitelském týmu a výpovědní hodnota výsledků takové studie bývá přesně definována. Při čtení výsledků menších klinických hodnocení je třeba brát v úvahu také jejich design a způsob statistického vyhodnocení. Obzvláště to platí pro výsledky nerandomizovaných menších souborů pacientů z klinické praxe. Možnost zavádějících závěrů zde bývá vysoká.

Literatura:

1. Dušek L.: Klinické studie v onkologii - analýza a management dat. V: Obecná onkologie a podpůrná léčba. Adam Z., Vorlíček J., Kostíková J. (eds), Grada 2003, 745-748
2. Green SB: Basics of Clinical Trial Design and Interpretation, American Society of Clinical Oncology, 1999 Educational Book
3. Green SB: Interpreting Clinical Trials: A Review, American Society of Clinical Oncology, 2003 Educational Book
4. Havránek T.: Statistika pro biologické a lékařské vědy, Academia Praha, 1993. 480s

5. Kasal P., Svačina Š.: Lékařská informatika, Praha: Karolinum, 1998. 543s
6. Libiger J.: Placebo: klamání nemocného nebo nástroj poznání? Psychiatrie 2003; 7(4):290-300.
7. Sargent D.J., Wieand S., Benedetti J., Labianca R., Halley D.G., Shepherd L.E., Seitz J.F., Francini G., De Gramont A., Goldberg R.M.: Disease free survival (DFS) vs. Overall survival (OS) as a primary endpoint for adjuvant colon cancer studies: Individual patient data from 12,915 patients on 15 randomized trials. Journal of Clinical Oncology, 2004 ASCO, vol. 22, 14S (July 15 Suppl.), 2004, 3502

HODNOCENÍ ODPOVĚDI NÁDOU NA LÉČBU - RECIST

RESPONSE EVALUATION CRITERIA IN SOLID TUMORS

DEMLOVÁ R

ODDĚLENÍ KLINICKÝCH HODNOCENÍ, MASARYKŮV ONKOLOGICKÝ ÚSTAV, BRNO

Souhrn: Cílem sdělení je objasnění pojmu RECIST kriteria, která jsou používána k hodnocení léčebné odpovědi v klinických studiích v onkologii a vymezení rozdílů mezi RECIST a WHO kritérii.

Klíčová slova: RECIST kriteria, WHO kriteria, měřitelné léze, neměřitelné léze, hodnotitelné léze, cílové léze, necílové léze

Summary: This article will explain the definition of RECIST criteria, commonly used in oncological clinical trials, and definition of differences between RECIST and WHO criteria.

Key words: RECIST criteria, WHO criteria, measurable lesions, nonmeasurable lesions, evaluable lesions, target lesions, non-target lesions

Hodnocení odpovědi nádoru na léčbu je nedílnou součástí vyhodnocování účinnosti léčby nejen v onkologických klinických studiích, ale je běžně používáno i v klinické praxi. Pojmy jako kompletní nebo parciální remise, stabilizace nebo progresse jsou užívány všemi lékaři, kteří se podílí na léčbě onkologických pacientů.

Jednou z prvních studií, která obecněji definovala určitá kriteria léčebné odpovědi, byla studie Zubroda et al. publikovaná v roce 1960. „Léčba je považována za efektivní, dojde-li k zmenšení celkové měřené hmotnosti nádoru za současné podmínky, že nedojde k progresi žádných z jednotlivých lézí a současně se neobjeví žádná léze nová“ (1).

Následně z iniciativy WHO (World Health Organisation) došlo v letech 1977 a 1979 k několika setkáním, na kterých byl přijat konsensus stran hodnocení odpovědi na léčbu, stejně tak jako k definování kritérií hodnocení nežádoucích účinků (2,3). Takto definovaná kriteria se pak dostala do obecného povědomí jako „WHO kriteria“ a byla dlouhou dobu užívána v široké klinické praxi. WHO kriteria byla postavena na dvou rozměrném měření nádorových ložisek (v nejdelší ose a na ní kolmém rozměru), což v praxi přinášelo poměrně obtížně precizní měření s velkou časovou náročností pro hodnotícího radiologa.

Koncem 90. let se pak aktivity řady kooperativních skupin v onkologii shodly na myšlence, že by byla potřeba revokovat diskusi o „měření nádorů“ a případně dosavadní kriteria pozměnit. Nové návrhy přišly s myšlenkou měření nádorového ložiska pouze v jednom nejdelším rozměru, což by alespoň částečně práci zjednodušilo, samozřejmě za současného předpokladu, že výsledky hodnocení budou odpovídat dosud používaným standardům.

Nově navržená kriteria, tzv. RECIST kriteria (Response Evaluation Criteria in Solid Tumors), pak byla aplikována ve 14

klinických studiích, kde byly u těchto pacientů současně použity i WHO kriteria s tím, že následně došlo k vyhodnocení případných neshod při porovnání léčebné odpovědi (4,5). Rozdílné odpovědi se objevily jen v zanedbatelném procentu a bylo tak potvrzeno, že nově navržená RECIST kriteria jsou srovnatelná s dosud používanými WHO kritérii. V čem se obě kriteria nejpodstatněji vzájemně liší ukazuje přehledně tabulka 1. WHO kriteria pracují s pojmy „měřitelné“ (measurable), „neměřitelné“ (non-measurable) a „hodnotitelné“ (evaluable) onemocnění. Jako „měřitelné“ se označují ty léze, které jsou měřeny ve dvou na sebe kolmých rozměrech, k jejich zobrazení je použita některá z reprodukovatelných radiodiagnostických metod s výjimkou sonografie, a jejich rozměry přesahují hodnotu 20 mm v jednom a 10 mm ve druhém rozměru. Všechny menší léze, popřípadě kostní metastázy, jsou označovány jako „neměřitelné“.

„Měřitelnost“	WHO kriteria	RECIST kriteria
Měřitelné léze	Léze měřeny ve dvou na sebe kolmých rozměrech*	Léze měřeny pouze v nejdelším rozměru (LD)**
Neměřitelné léze	Všechny ostatní menší léze	Všechny ostatní včetně lézí malých rozměrů
Hodnotitelné léze	Např. ascites, fluido thorax, peritoneální rozsev...	Termín není používán
„Odpověď“	WHO kriteria	RECIST kriteria
Kompletní remise (CR)	Vymizení všech lézí	
Parciální remise (PR)	≤ 50 % ve srovnání se vstupním vyšetřením	≤ 30 % ve srovnání se vstupním vyšetřením
Stabilizace (SD/non-PD)	Nesplnění kritérií pro PD, PR nebo CR	Nesplnění kritérií pro PD, PR nebo CR
Progrese (PD)	> 25 % ve srovnání se vstupním vyšetřením	> 20 % ve srovnání se vstupním vyšetřením

Tabulka č. 1

*/ rozměry musí být minimálně 20 mm v jednom a 10 mm ve druhém rozměru

**/ LD : longest diameter (jako měřitelné mohou být označeny pouze v případě, že LD ≥ 20 mm při použití konvenčních zobrazovacích metod nebo LD ≥ 10 mm při použití spirálního CT)

Tabulka č. 2: Recist kriteria

Cílové léze (Target lesion) ¹		
Kompletní remise (odpověď)	CR	Vymizení všech cílových lézí
Parciální remise (odpověď)	PR	Minimálně 30% pokles součtu nejdelších průměru (LD) ve srovnání s úvodním měřením
Progrese	PD	Minimálně 20% nárůst součtu nejdelších průměru (LD) nebo objevení se 1 či více nových lézí
Stabilizace onemocnění	SD	Nesplnění kritérií pro PD, PR nebo CR
Ne-cílové léze (Non-target lesions) ²		
Kompletní remise (odpověď)	CR	Vymizení všech ne-cílových lézí a normalizace nádorových markerů
Nekompletní remise/stabilizace	Non-CR/SD	Přetrvávání jedné nebo více necílových lézí a/nebo přetrvávání hladin nádorových markerů nad normu
Progrese	PD	Jednoznačná progrese existujících necílových lézí nebo objevení se 1 či více nových lézí

Celková terapeutická odpověď

Cílové léze	Ne-cílové léze	Nové léze	Celková odpověď
CR	CR	Ne	CR
CR	non-CR/SD	Ne	PR
PR	non-PD	Ne	PR
SD	non-PD	Ne	SD
SD	jakákoliv odpověď	Ano nebo Ne	PD
jakákoliv odpověď		Ano nebo Ne	PD
jakákoliv odpověď	jakákoliv odpověď	Ano	PD

RECIST kriteria :

Celková terapeutická odpověď:

¹Target lesions: měřitelné léze, definovány jako měřitelné v jednom nejdelším rozměru a to ≥ 20 mm při použití konvenčních zobrazovacích technik (konvenční CT, RTG, MRI, PET) nebo ≥ 10 mm při použití spirálního CT

²Non-target lesions: neměřitelné léze, definovány jako všechny ostatní léze <20 mm při konvenčním zobrazení nebo <10 mm při použití spirálního CT. Dále kostní léze, leptomeningeální nádory, ascites, pleurální/periakardiální výpotek, lymphangitis, zánětlivé karcinomy prsu a cystické léze jsou rovněž

Hodnotitelné jsou pak příkladem léze typu ascitu, fluidothoraxu, perikardiálního výpotku nebo rozsev na peritoneu. RECIST kriteria pak rozlišují pouze „měřitelné“ a „neměřitelné“ léze s následným rozdělením těchto ložisek na tzv. „cílové“ léze (target lesions) a „necílové“ (non-target). RECIST kriteria neppracují s pojmem „hodnotitelné“.

Podstatným rozdílem je především způsob měření, kdy na rozdíl od WHO kritérií RECIST kriteria vyžadují měření pouze v jednom nejdelším rozměru. Následně se pak liší také definice parciální remise a progrese onemocnění. O parciální remisi hovoříme v případě zmenšení sumy jednotlivých ploch měřitelných lézí o více než 50 % oproti vstupnímu měření u WHO kritérií, respektive zmenšení o více než 30% součtu nejdelších lézí u RECIST kritérií. V případě progrese se jedná o navýšení o 25% u WHO kritérií, respektive 20% u RECIST.

RECIST kriteria (tab. 2) se v klinických studiích používají nyní již téměř standardně, jejich výhodou je především snadnější „měřitelnost“ z důvodu nutnosti zaznamenat pouze jeden nejdelší rozměr sledované léze, přičemž reprodukovatelnost a způsob vyhodnocení vychází z podobných principů jako tomu bylo u WHO kritérií.

Literatura :

1. Zubrod CG, Schneiderman M, Frei E III, Brindley C, Gold GL, Shnider B et al. Appraisal of methods for the study of chemotherapy of cancer in man: comparative therapeutic trial of nitrogen mustard and triethylen thiophosphoramidate. J Chron Dis, 11: 7-33
2. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. Cancer 1981, 47:207-214
3. Green S, Weiss GR. Southwest Oncology Group standard response cri-

teria, endpoint definitions and toxicity criteria. Invest Newe Drug 1992, 10:239-253

4. Therasse P, Arbuck SG, Eisenhauer AE, Wanders J, Kaplan RS, Rubinstein L et al. New Guidelines to evaluate the response to treatment in solid tumors. J Natl Cancer Inst 2000, 92:205-216
5. James K, Eisenhauer E, Christian M, Terenziani M, Vena D, Muldal A et al. Measuring of response in solid tumors: unidimensional versus bidimensional measurements. J Natl Cancer Inst 1999, 91:494-495

VÝBĚR PACIENTŮ V KLINICKÝCH STUDIÍCH - RIZIKA PRO INTERPRETACI V BĚŽNÉ KLINICKE PRAXI

SELECTION OF PATIENTS IN CLINICAL TRIALS - RISKS FOR INTERPRETATION IN ROUTINE PRACTICE

DEMLOVÁ R.

ODDĚLENÍ KLINICKÝCH HODNOCENÍ, MASARYKŮV ONKOLOGICKÝ ÚSTAV, BRNO

Souhrn: Každý pacient může vstoupit do klinické studie pouze v případě splnění všech vstupních kritérií daných protokolem studie. Lze tedy zjednodušeně říci, že populace studiových pacientů je populací výběrovou. Cílem tohoto sdělení je zamýšlení se nad možnými riziky při interpretaci výsledků studií v běžné klinické praxi, respektive v širší populaci běžně léčených onkologických pacientů.

Klíčová slova: vstupní kritéria, vylučující kritéria, studiová populace pacientů, správná klinická praxe

Summary: The study treatment should be offered only to those patients, which meet all inclusion criteria in accordance with study protocol. We can say the study population is more or less selected. This article will mention eventual problems with interpretation of the clinical study results in common practice.

Key words: inclusion criteria, exclusion criteria, patient study population, good clinical practice

Současné standardy onkologické léčby se v době „medicíny založené na důkazech“ opírají o výsledky randomizovaných klinických studií a stejně tak opačně, pozitivní výsledky správně naplánovaných, kontrolovaných multicentrických studií mohou změnit dosavadní standardy v léčbě konkrétních nádorových onemocnění.

Každý protokol klinické studie, je-li plánován podle zásad tzv. správné klinické praxe neboli Good Clinical Practice (GCP), má své zcela standardní části a nepodkročitelné kapitoly. Jednou ze základních součástí protokolu je výčet tzv. „inclusion a exclusion“ kritérií, podle kterých jsou pacienti do studie vybíráni¹. Každý z potenciálních subjektů hodnocení, neboli jednoduše řečeno vhodných pacientů, musí naprosto jednoznačně vyhovět všem kritériím daným protokolem studie. Nesplnění i pouze jednoho ze vstupních nebo vylučujících kritérií je důvodem pro nezařazení pacienta do klinické studie. Stejně tak protokol studie definuje velmi jednoznačně i průběh studie z hlediska podávané léčby včetně případných modifikací dávkovacích schémat, jednotlivých kontrolních vyšetření i následného sledování po ukončení aktivní léčby.

Každé nedodržení protokolu studie je považováno za tzv. odchylku od protokolu a v případě závažnějších chyb, tzv. „protocol deviation“, se může stát, že data o výsledcích léčby tohoto pacienta nejsou zahrnuta do závěrečné analýzy. Je nesporné, že takto naplánovaná a kontrolovaná klinická studie je vědecky zdůvodnitelná a výsledky srovnání kontrolní a experimentální větve studie pak mohou vést k relevantním závěrům, neboť tzv. srovnávají srovnatelné.

Aby bylo možno výsledky klinické studie správně interpretovat, je nutno posuzovat zkoumaný soubor právě z hlediska správnosti zařazování pacientů a rozložení souboru z hlediska jednotlivých stratifikačních parametrů. Jako příklad ne zcela správně rozloženého souboru pacientů v jednotlivých srovnávaných ramenech může být studie finských autorů srovnávající adjuvantní podávání klodronátu u pacientek s kar-

cinomem prsu a pozitivními uzlinami, které byly primárně adjuvantně léčeny chemo- nebo hormonoterapií². Současně byly pacientky randomizovány do dvou větví studie, a to s nebo bez adjuvantního podání klodronátu v dávce 1600 mg po dobu tří let. Analýza po deseti letech sledování došla k poměrně překvapivým výsledkům, kdy obě skupiny pacientek byly zcela srovnatelné z hlediska četnosti výskytu kostních metastáz, ve skupině pacientek s klodronátem však došlo k zvýšenému výskytu non- skeletálních rekurencí. DFS (disease free survival) byl u ER+ pacientek v obou ramenech léčby bez rozdílu, u ER negativní skupiny byl však ve skupině s klodronátem signifikantně kratší. Při podrobnějším rozboru výsledků je však nutno poukázat na zcela zásadní skutečnost, že procento ER a PR negativních pacientek bylo ve skupině léčené klodronátem signifikantně vyšší než u skupiny kontrolní. Postmenopauzální pacientky s negativními receptory byly navíc léčeny hormono-, nikoliv chemoterapií. Všechny tyto skutečnosti významně ovlivnily přežití ve srovnávaných skupinách a výsledky pak jistě mohou být takto kontroverzní.

Nutno poznamenat, že takto nesprávně stratifikované soubory se ve velkých multicentrických studiích objevují zcela minimálně a většina studií tak může vést ke zcela obhajitelným závěrům.

Při hlubším zamýšlení však může nastat problém při zavedení takto získaných optimistických „statisticky signifikantních“ závěrů klinické studie do rutinní klinické praxe. Nová léčba je následně nabídnuta i pacientům, kteří by primární přísná vstupní a vylučující kritéria rozhodně nesplnili (zejména z hlediska parametrů případných komorbidit, konkomitanti medikace nebo laboratorních parametrů určujících případné orgánové rezervy).

Mengis a kol.³ analyzovali ve své práci léčebné výsledky pacientů s akutní myeloidní leukémií, kteří byli léčeni buď v rámci protokolu klinické studie nebo mimo něj, avšak stejným léčebným schématem. Pacienti léčeni mimo klinickou studii

dosáhli horších léčebných výsledků, které byly primárně ovlivněny právě výběrem, respektive správněji spíše „nevýběrem“ pacientů. Mimo protokol se ocitli pacienti, kteří nesplnili vstupní kritéria, především z hlediska horní věkové hranice nebo případných závažnějších komorbidit.

Do klinických studií mohou být velmi problematicky zařazováni také pacienti, kteří mají rychle progredující onemocnění nebo se nacházejí v život ohrožující situaci a zahájení chemoterapie nebo jiné léčby nelze v žádném případě odkládat (např. pacienti s SCLC a obstrukcí horní duté žíly nebo rychle progredující high-grade lymfomy). Problémem pro jejich zařazení je kompletace nutných vstupních tzv. baseline vyšetření, bez kterých nemohou být do studie zařazeni a jejichž naplánování vyžaduje určitou nezbytnou dobu. Tito pacienti, představující většinou prognosticky závažnou kategorii, jsou z participace v klinických studiích předem vyloučeni a takto nepřímo selektují „prognosticky příznivější“ skupinu pacientů do klinických studií zařazovaných.

Stejně tak problematicky se ve světle nejlepších stávajících diagnostických metod může v onkologii jevit tzv. „staging“ nádorového onemocnění a na základě jeho určení léčba v rámci dané klinické studie. Příkladem, klasifikace „III klinické stadium“ při zahájení studie neznamená, že v budoucnu se bude jednat o zcela stejně rizikovou skupinu pacientů, použijeme-li případné nejnovější diagnostické metody - v literatuře popisovaný tzv. Will Rogers fenomén „migrace stadií“⁴.

Z výše uvedených faktů vyplývá, že onkologové jsou tímto postaveni před poměrně závažné dilema. Klinický výzkum plně respektuje exaktní výzkumné standardy s přísnými kontrolními mechanismy, jež jsou akceptovatelné pro regulatorní i vědecké autority. Výsledky takto prováděného klinického výzkumu mohou pozměnit stávající léčebné standardy. V okamžiku uvedení těchto původně výzkumných léčebných sché-

mat do rutinní klinické praxe jsou však veškerá „výběrová kritéria“ referenční studie včetně dalších podmínek studie jako by zapomenuta a nikdo již neřeší, jak může tato skutečnost ovlivnit skutečné reálné léčebné výsledky.

Tuto skutečnost si řada onkologů plně uvědomuje a existují kooperativní skupiny, jež plánují studie s tzv. „open-mind design“. Jako příklad je možno uvést „International Adjuvant Lung Cancer Trial“⁵, jež povoluje 148 participujícím centřům zařadit pacienty podle individuálně hodnoceného klinického stadia, dávky cisplatinu i výběru léčiva v kombinaci. Jak lze předpokládat, studie je kritizována právě pro zcela volná kritéria a heterogenitu léčby, autoři však argumentují „reálnějšími výsledky“ napříč centry a tím i zobecněním léčebných výsledků.

Stejný příklad poskytuje International Collaborative Ovarian Neoplasm (ICON) 3 Trial⁶ s indikací paclitaxelu v první linii léčby, kde se vstupní kritérium omezuje na sdělení „klinik si je jist indikací první linie a vhodnosti pacienta k léčbě“. O tom, jak zcela odlišný je to přístup ve srovnání s klasicky plánovanými klinickými studiemi, asi netřeba polemizovat.

Jaké tedy může být rozřešení tohoto stávajícího problému? Jistě ne jednoduché, nicméně určitá úvaha by zde byla. Je zcela bez pochyby, že studie fáze I a II, jejichž cílem je detailní zhodnocení antitumorozního efektu, nalezení optimálního dávkovacího schématu a hodnocení případných nežádoucích účinků, se neobejde bez přísných kritérií výběru pacientů a detailně zpracovaného protokolu studie. U velkých randomizovaných studií fáze III, srovnávajících standardní léčbu s novými léčebnými přístupy, by však pravděpodobně bylo možno přistoupit k méně přísným vstupním kritériím, případně k výše zmíněnému „open-mind design“. Tato strategie by mohla lépe reflektovat následnou situaci v následné každodenní léčebné praxi^{7,8}.

Literatura :

1. Martin F Fey, Selection of the patients, The Lancet Oncology Vol 5, December 2004
2. Saarto T. et al., Ten-year follow-up of randomised controlled trial of adjuvant clodronate treatment in node-positive breast cancer patients
3. Mengis C, Aebi S, Tobler A et al., Assessment of difference in patient populations selected for or excluded from participation in clinical trial phase III acute myelogenous leukemia trials. J Clin Oncol 2003, 21:3933-3939
4. Feinstein AR, Sosin DM, Wells CK. The Will Rogers Phenomenon. Stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. New Eng J Med 1985, 312:1604-1608

5. Arriagada R, Bergam B, Dunant A et al. Cisplatin based adjuvant chemotherapy in patients with completely resected NSCLC. New Eng J Med 2004, 350:351-360
6. International Collaborative Ovarian Neoplasm (ICON) Group. Paclitaxel plus carboplatin versus standard chemotherapy with either single-agent carboplatin or cyclophosphamide, doxorubicin and cisplatin in woman with ovarian cancer: the ICON3 randomised trial. Lancet 2002, 360:2086
7. Yusuf S, Collins R, Peto R. Why we need some large, simple randomised trials ? Statistics Med 1984, 3:409-420
8. Holmberg L, Baum M. Can results from clinical trials be generalised ? Nature Med 1995, 1:734-736

METODICKÉ PROBLÉMY KLINICKÝCH STUDIÍ V ONKOLOGII

METHODOLOGICAL PROBLEMS OF CLINICAL TRIALS IN ONCOLOGY

NOVOTNÝ J.¹, PETRUŽELKA L.¹, PECEN L.²

¹ONKOLOGICKÁ KLINIKA VFN A 1. LF UK, PRAHA

²ÚSTAV INFORMATIKY AV ČR, PRAHA

Souhrn: Randomizované klinické studie jsou v dnešní době považovány za nejhodnotnější vědecký důkaz, který má ve svém konečném důsledku nalézt nový léčebný postup pro vybranou populaci pacientů, který je ve srovnání s dosavadním postupem účinnější, bezpečnější, levnější, anebo má více těchto pozitivních charakteristik současně. V článku se zabýváme problémy, které se objevují při využívání takových studií jako klíče při rozhodování o léčebném postupu u konkrétního pacienta.

Klíčová slova: randomizovaná klinická studie, metaanalýza, metodické otázky

Summary: Randomized clinical trials are considered to be one of the most important tools, that allow direct comparison of current standard and experimental therapies. Subsequently, the more efficient and/or better tolerated regimen is selected as the new standard. This article focuses on the problems with the interpretation of trial results as well as on the use of these trials in the selection of appropriate therapies for the individual patient.

Key words: randomized clinical trial, metaanalysis, methodological issues

Úvod

Randomizované klinické studie jsou stále častěji považovány za nejhodnotnější vědecký důkaz, který má přinést klinické praxi odpověď na otázku, zda nová experimentální léčba „B“ přináší našim nemocným více než stávající standardní léčba „A“.

Pokusme se ale zamyslet nad tím, zda tuto jasnou odpověď můžeme od klinických studií očekávat.

Problémy s designem studií

V tomto článku nechceme diskutovat otázky spojené s úvodními fázemi klinických zkoušek, ale zaměřujeme se na výsledky studií třetí fáze klinických zkoušek, které mají poskytovat údaje, jež umožňují provádění správné lékařské praxe. K tomuto účelu nejlépe slouží rozsáhlé randomizované studie¹⁾, ve kterých se jeden nebo více postupů experimentálních srovnává s ramenem standardním. V případě příznivého výsledku pak experimentální rameno nahrazuje stávající standard a stává se tak standardem novým. V randomizované studii můžeme testovat nejen nový léčebný postup, ale také i standardní léčebné postupy v kombinaci s potenciálním prognostickým či prediktivním faktorem. Randomizovaně může být též potenciální prognostický či prediktivní faktor stanovován. Vždy ale v randomizované studii musí vystupovat nějaký léčebný postup buď i standardní²⁾.

Některé typy designů randomizovaných studií užívaných v onkologii ukazuje schématicky obrázek 1 (a to v souvislosti s prediktivním faktorem/markerem).

V onkologii se až na výjimky u randomizovaných studií fáze I obvykle nepoužívá crossover design, kde jsou v randomizovaném pořadí aplikovány dva nebo více léčebných postupů. Bohužel, ne vždy jsou randomizované studie připraveny důkladně, a tak se nezdá setkání s problematickým návrhem studie, který je krátce po její publikaci důvodem nespokojenosti řady kliniků. Jednou z nejčastějších chyb designu je nesprávně zvolené srovnávací rameno. Z nejméně tak-

to kritizovaných studií zmiňme např. studie hodnotící účinnost adjuvantní hormonoterapie proti chemoterapii, ve kterých se opakovaně setkáváme s neadekvátním chemoterapeutickým režimem, zařazováním pacientek s negativními hormonálními receptory v primárním nádoru, s nepodáním hormonální léčby pacientkám léčeným chemoterapií apod. Do této kategorie patří také studie s paklitaxelem v adjuvantní léčbě karcinomu prsu, které neodpovídají na otázku, zda je lepší léčit paklitaxelem či jen atracykliny, ale pouze poukazují na nedostatečnost 4 cyklů chemoterapie AC pro pacientky s uzlinovým postižením. Zatímco v některých případech výše uvedených studií musíme být shovívaví, neboť to, co se z dnešního pohledu zdá nesprávné, nebylo v době zahajování studií známo, jindy je volba kontrolního ramene zcela nepochopitelná. Příkladem takovéto prvotně nesprávné volby je „standardní“ režim paliativní chemoterapie karcinomu žaludku cisplatin - 5-fluorouracil, který byl testován proti experimentálnímu ramenu docetaxel - cisplatin - 5-fluorouracil. V literatuře však nenajdeme jedinou studii, která by nám dovolila zmíněný režim cisplatin - 5-fluorouracil považovat za standard paliativní léčby nádorů žaludku. Podobných příkladů bychom mohli uvést celou řadu.

Druhou běžnou chybou randomizovaných studií je nedostatečná velikost souboru nemocných. Ta bývá způsobena nesprávnou statistickou konzultací, nebo, což je mnohem častější, nedostatečným nábořem pacientů do studie. Bylo zjištěno, že třetina studií prováděných v USA se účastní méně než 75% plánovaného množství pacientů³⁾. Proto řada srovnání, která nejsou statisticky významná, může být způsobena právě nedostatečnou možností tyto rozdíly postihnout⁴⁾. Typickým příkladem tohoto jevu je celá plejáda studií, která zkoumala experimentální dvojkombinace léčiv proti monoterapii gemcitabinem u nemocných s inoperabilním nebo lokálně pokročilým karcinomem pankreatu, a která až na dvě výjimky neprokázala žádný přínos kombinované léčby na celkové přežívání nemocných. Minimální efekt, který dvojkombinace mohou u tohoto c-

hemorezistentního onemocnění přinést, by totiž znamenal nutnost extrémně velkých studií a pak, musíme také zvážit, zda je statisticky prokázán efekt také klinicky relevantní (např. prodloužení přežívání nákladnou léčbou o 1-2 týdny).

Problémy s výběrem nemocných

Velkou chybou mnoha studií je, že z nich jsou často vyřazováni starší nemocní nebo nemocní s alterací jaterních či ledvinných funkcí. Tento zvyk pochází z klinických testů léků na méně závažná onemocnění, ale tento postup není v případě onkologických onemocnění vždy optimální, protože právě podíl těchto prognosticky méně příznivých pacientů není v rutinní klinické praxi zanedbatelný. U nich se pak drobné zlepšení léčebných výsledků dosažené experimentálním postupem „B“ nemusí vůbec projevit. V případě kaskády předregistračních studií sice musí studie zaměřená na starší nemocné či na nemocné s alterací jaterních či ledvinných funkcí nutně existovat, může být ale rozsahem mnohem menší, ačkoliv v pozdější rutinní aplikaci nového léku je podíl takových pacientů značný.

Problémy s interpretací výsledků

Při čtení závěrů studie je klinik zavalen velkým objemem dat, jejichž interpretace bez dostatečné znalosti statistiky může být zavádějící.

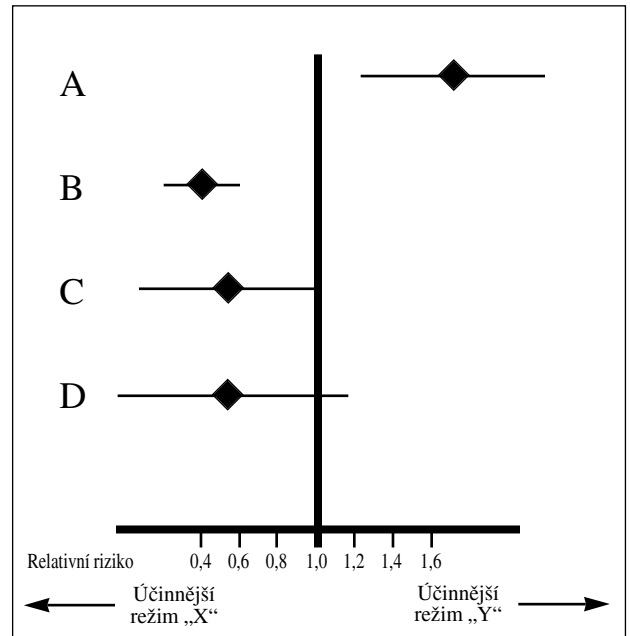
Máme-li k dispozici publikaci plánované průběžné analýzy, jde zcela určitě o závěr hodný zřetele. Průběžné analýzy se totiž publikují především tehdy, vyskytnou-li se nějaké neočekávané nežádoucí účinky léčby, o kterých by měl být klinik včas informován, anebo dochází k významnému zlepšení celkového přežívání nebo kvality života nemocných. V poslední době je však předčasně ukončování studií na základě plánovaných analýz poměrně časté i z jiných důvodů. To je však značně kontraproduktivní⁵⁾. Z nedávno takto ukončených studií zmiňme SWOG 9701 (udržovací léčba paklitaxelem 12 vs 3 měsíce) a MA-17 (5-leté užívání letrozolu vs placebo po 5 letech tamoxifenu). O tom, zda má být studie předčasně ukončena, zda má být studie odslepena a pacientům s méně účinnou nebo rizikovější léčbou má být nabídnuta alternativa, bdí nezávislá „Data and Safety Monitoring Board“. Důvodem k ukončení studie by rozhodně nemělo být zlepšení času do progresu do onemocnění, neboť v mnoha dříve provedených studiích bylo prokázáno prodloužení času do progresu nemocí (čas bez známek onemocnění), aniž by bylo jakkoliv ovlivněno celkové přežívání (např. ve studii GOG 132 a dalších). Tento rozpor je dán především existencí účinné léčby druhé a další linie, která umožňuje u relabujících pacientů opět získat kontrolu nad nádorovou nemocí. A tak v době, kdy se snažíme praktikovat medicínu založenou na důkazech, by neměla být standardem pouhá víra ve skutečnost, že prodloužení času do progresu nebo času bez známek onemocnění bude jistě doprovázeno i prodloužením přežívání. Z takového stavu nemůže mít prospěch ani pacient s lékařem, kteří nevědí, jak postupovat co neefektivněji, tak ani výrobce léčiva, který s neúplnými daty obtížně přesvědčí váhající lékaře o správnosti nasazení svého produktu.

Čteme-li později kompletní sdělení o randomizované studii, musíme velmi opatrně interpretovat závěry studií, abychom zjistili, zda je výsledek statisticky významný či nikoliv. V případě, že je studie postavena jako tzv. non-inferiority studie (např. kapecitabin vs FUFU Mayo v adjuvantní léčbě karcinomu tračníku), pak nesignifikantní p-value naznačují horší účinnost experimentálního postupu, zatímco signifikantní p-value (tj. pod 0,05 při tradičně užívané hladině významnosti či 0,025 jak to doporučují FDA guidelines v případě jednostranných testů) ukazují, že experimentální postup je „srovnatelný či lepší“ než standardní. Z takto postavené studie nejsme nikdy schopni prokázat přednost ramene experimentálního před standardním. Naproti tomu při designu studujícím superioritu režimu A před B nemůžeme nikdy prokázat, že experimentální postup je horší než standardní.

Poměrně často se také setkáváme s pojmem nesignifikantní tr-

end. Za tímto pojmem se skrývá zcela nesignifikantní diference, tj. diference která nebyla zatím nijak statisticky prokázána a nemá tedy žádný statistický význam. Ale zvláště v prezentacích farmaceutických firem se tento trend používá k přesvědčení lékařů o vhodnosti právě jejich produktu.

V případě použití konfidenčních intervalů pro poměr hazardů mezi rameny A a B je třeba upozornit na skutečnost, že pokud konfidenční interval obsahuje hodnotu/bod 1,00 je to vždy známkou nesignifikantního výsledku, byť je střed uvedeného konfidenčního intervalu pro poměr hazardů sebevše vychýlen na jednu či druhou stranu.



Obrázek 2.

Pozornost si jistě zaslouží také referování relativních rizik a poměru hazardů. Relativní snížení rizika recidiv o polovinu je sice zajímavý údaj, ale kolik nemocných se opravdu vyvaruje lepším postupem recidivy? Relativní riziko 0,5 v sobě může skrývat jak vysoce efektivní postup - např. redukcí recidiv 30% vs 60%, tak i něco velmi nepatrného (např. rozdíl 0,003% vs 0,006%). Proto je třeba vždy zkontrolovat absolutní rozdíly sledovaných parametrů mezi jednotlivými rameny. Aby problém nebylo málo, hned narážíme na další, který je ale bohudíky moderními statistickými postupy již vyřešen. Je-li primárním cílem např. čas bez známek recidivy, čas do lokálního selhání nebo čas bez známek progresu onemocnění, je třeba disponovat statistickým nástrojem, který eliminuje chybu vzniklou tím, že nemocný např. podlehne jiné souběžné nemoci nebo recidivuje ve vzdálených lokalizacích, zatímco my hodnotíme lokální kontrolu (např. u radioterapie). Ne vždy je ale optimální statistický postup skutečně použit a čtenář-lékař to ani nemusí poznat.

Za určitých okolností bývají publikovány také údaje o předmě neplánovaných analýzách. Ty by podle zásad správné klinické praxe neměly být vůbec zveřejňovány, neboť právě proto, že nejsou předem plánovány, mohou obsahovat některé potenciálně zavádějící údaje - např. heterogenitu ve věku, pohlaví, charakteristikách nemoci, výkonnostním stavu apod. Výsledky těchto analýz je tedy třeba považovat jen za orientační.

Často řeší jeden problém velké množství klinických studií - např. chemoterapii kolorektálního karcinomu, radioterapii karcinomu prostaty apod. Proto vznikají systematická review, tzv. metaanalýzy, která se snaží zmenšit heterogenitu mezi jednotlivými studiemi a učinit závěry lépe využitelné v klinické praxi. V zásadě se setkáváme se dvěma typy metaana-

lýz. Metaanalýza typu „literature search“ využívá k získání závěru definic studovaných populací (včetně přesných inclusion/exclusion kritérií) hodnot p-value, konfidenčních intervalů a dalších proměnných, které jsou publikovány ve sbornících abstrakt, na přednáškách či v odborných časopisech. Ovšem hodnotnější a zcela nejvýše na pomyslném žebříčku kvality poskytovaných údajů stojí metaanalýza prováděná ze zdrojových dat. Při práci na této analýze oslovuje statistický tým autory jednotlivých studií, sjednocuje metodiku statistického zpracování a získává aktualizovaná data o přežívání, počtu léčebných odpovědí apod. Ne vždy jsou ale autoři jednotlivých studií ochotni spolupracovat na této metaanalýze prováděné ze zdrojových dat.

Pokud se tedy rozhodujeme o volbě léčebného postupu na základě medicíny založené na důkazech, pak nejprve zohledníme výsledky metanalýz, poté rozsáhlých randomizovaných studií fáze III, a až nakonec studií fáze II. Pokud neexistuje standardní postup, zvažujeme zařazení nemocných do klinických studií, jsou-li k dispozici.

Problémy s publikací výsledků

Při výčtu problémů, které se vyskytují při práci s klinickými studiemi, nelze opomenout ani tak zdánlivě jednoduché téma, kterým je publikace výsledků.

Bylo zjištěno, že větší množství studií s pozitivními výsledky je otiskováno v anglicky psané literatuře, a naproti tomu řada hodnotných prací z evropské či východoasijské oblasti je otištěna pouze v mateřském jazyce v případě, že experimentální postup není účinnější než postup standardní.

Také skutečnost, zda studie byla publikována pouze formou abstraktu nebo v plném znění je zavádějícím faktorem. Nedávno zveřejněná analýza randomizovaných studií v pediatrii prokázala, že pouze 59,1% abstraktů bylo následně publikováno *in extenso*. Častěji byly kompletně publikovány studie s větším počtem

nemocných, s kvalitnějším designem a s pozitivními výsledky⁶⁾. Takto může nepublikování některých podstatných zjištění z mnoha klinických studií vést ke snížení kvality léčebně-preventivních rozhodnutí pro dané skupiny nemocných, což ve svém konečném důsledku může ohrozit život konkrétního pacienta.

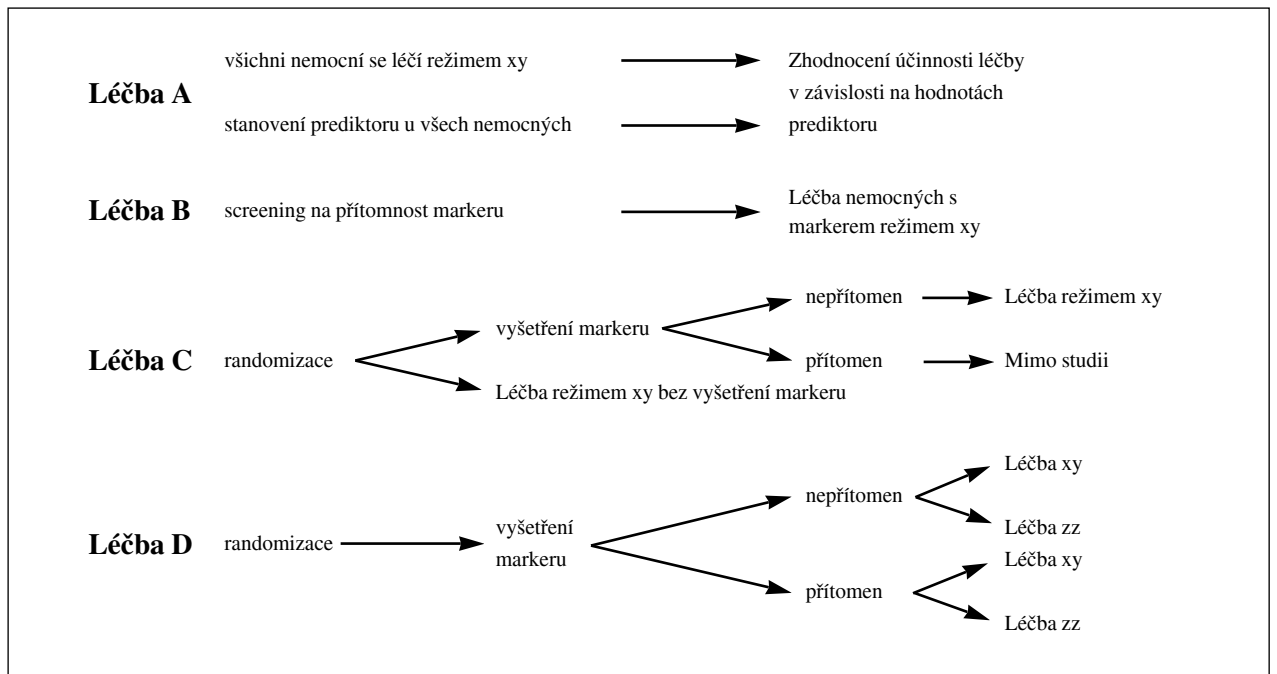
Větší pozornost by také měla být věnována přesnosti citací při přípravě publikací. Jejich kontrola provedená ve čtyřech velkých otorhynolaryngologických časopisech ukázala na přítomnost chyb u 37,5% všech citací; chyby hodnocené jako závažné byly odhaleny u 11,9% všech citací⁷⁾. S nesprávnými, a někdy i neexistujícími citacemi, se často setkáváme v propagačních materiálech farmaceutických firem.

V poslední době se lze setkat v publikační praxi nejprestižnějších časopisů s tzv. „zrychleným uveřejněním klíčových sdělení“. V podstatě se jedná o přednostní uveřejnění studie, kterou autor, editor nebo recenzenti považují za velmi významnou a mající dopad na každodenní klinickou praxi. Problém této politiky spočívá v tom, jak vlastně hodnotit, která práce si zrychlené vydání zaslouží, a která nikoliv⁸⁾. Ghalli publikoval v roce 2002 výsledky průzkumu mezi lékaři, který ukázal, že do určité míry opravdu mají přednostně vydávané články vyšší váhu, na druhou stranu 5 z 12 kontrolních článků vydaných standardním způsobem mělo v očích lékařů vyšší hodnocení než články „urychleně“ vydané.

Závěr

Klinické studie prováděné v posledním desetiletí umožnily ustanovit nové, efektivnější standardy protinádorové léčby, které vedou k vyššímu počtu vyléčených pacientů. Současně ale také vyvolávají nové otázky, na které dnes neznáme odpověď, a s ohledem na biologii nádorových procesů ani dlouho znát nebudeme.

Nezbývá nám tedy, než při léčbě našich pacientů zohlednit kromě studií také vlastní zkušenost a intuici.



Obrázek 1.: Ukázky některých designu studií, které hodnotí klinickou využitelnost prediktivního markeru. Schéma A dokumentuje léčbu všech pacientů jedním režimem se simultánním stanovením prediktoru a zhodnocením výsledků léčby v závislosti na hodnotách prediktoru. Pokud známe z historických souborů účinnost nepredikované léčby „xy“, pak při designu B můžeme nalézt vyšší účinnost léčby „xy“, pokud ji podáváme pouze pacientům s potenciálně pozitivním prediktivním faktorem. Samozřejmě, že je zde silná možnost zkreslení (bias) rozdíly mezi zkoumanou populací a historickými soubory (ve stádiu onemocnění, věku, zastoupení pohlaví a mnoha a mnoha dalších faktorech včetně různých vedlejších onemocnění). Design C umožňuje v randomizovaném porovnání zjistit, zda prediktivní test zlepšuje ve srovnání s neselektovanou populací léčebné výsledky postupem „xy“. Ve studii s designem D hledáme účinnější režim pro nemocné se studovaným znakem, resp. bez něj.

Literatura

1. Altman D.G., Schulz K.F., Moher D., et al.: The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Int Med* 134, 2001, 663-694
2. Sullivan R.: Clinical trial design in oncology: protocol design. *Lancet Oncology* 5, 2004, 759-763
3. Charlson M.E., Horwitz R.I.: Applying results of randomized trials to clinical practice: impact of losses before randomisation *BMJ* 289, 1984, 1281-1284
4. Costa L.J., Xavier A.C., de Giglio A.: Negative results in cancer clinical trials - equivalence or poor accrual? *Controlled Clinical Trials* 25, 2004, 525-533.
5. *JCO* 2004, 22, 1542-1545
6. Klassen TP, Wiebe N, Russell K, et al.: Abstracts of randomized controlled trials presented at the society for pediatric research meeting: an example of publication bias. *Arch Pediatr Adol Med* 156, 2002, 474-479
7. Fenton J.E., Brazier H., De Souza A., et al.: The accuracy of citation and quotation in otolaryngology/head and neck journals. *Clin Otolaryngol Allied Sci* 25, 2000, 40-44
8. Ghali W.A., Cornuz J., McAlister F.A., et al.: Accelerated publication versus usual publication in 2 leading medical journals. *CMAJ* 166, 2002, 1137-1143

ODRAZ VÝSLEDKŮ KLINICKÝCH STUDIÍ V POPULAČNÍM ONKOLOGICKÉM REGISTRU

REFLECTION OF RESULTS FROM CLINICAL TRIALS IN POPULATION CANCER REGISTRY

ŽALOUĐÍK J., VYZULA R.

MASARYKŮV ONKOLOGICKÝ ÚSTAV A UNIVERZITNÍ ONKOLOGICKÉ CENTRUM BRNO

Souhrn: Randomizovaná klinická studie je sice dosud neúčinnějším vědeckým nástrojem klinického bádání, hledajícího optimální léčebné postupy, neměla by však zůstat jeho konečným cílem.

Rozhodující je, nakolik se podaří implementovat pozitivní výsledky klinických studií v rutinní praxi a jak se podaří prokázat a zopakovat také v neselektovaných souborech. V České republice lze základní dopady v praxi akceptovaných mezinárodních studií sledovat v Národním onkologickém registru vedeném od roku 1977, který obsahuje také diagnostické a léčebné parametry, a je vybaven analytickým softwarem SVOD (Systém pro vizualizaci onkologických dat), který vyvinulo Centrum biostatistiky a analýz MU v Brně. Rozdíly mezi výsledky klinických studií a praxí, zejména pak rozdíly mezi výsledky různých pracovišť zasahujícími do léčby, je možno hodnotit také z pohledu ekvity onkologické péče. V práci jsou uvedeny příklady, kdy nálezy v populačním onkologickém registru jsou v souladu s doporučeními z kontrolovaných klinických studií a jejich metaanalýz. Jde o adjuvantní chemoterapii u II. stadia karcinomu prsu a III. stadia karcinomu kolorekta, dále předoperační radioterapii karcinomu rekta, chemoadjuvanci ve stadiu III. maligního melanomu, srovnání léčebných metod u hepatocelárního karcinomu a vliv chemoterapie na vývoj III.-IV. stadií sarkomů měkkých tkání.

Klíčová slova: výsledky kontrolovaných klinických studií, národní onkologický registr, srovnání křivek přežití, medicína založená na důkazech, problém ekvity

Summary: Randomized clinical study represents to date the most effective measure for clinical research aimed to search for optimal therapies, but it should not remain its only goal. It is most important that the results of a controlled clinical study can be implemented in routine practice and repeated in non-selected groups of patients. In the Czech Republic the possibility exists to compare the impact of widely accepted protocols from clinical studies in nationwide population cancer registry managed since 1977. The parameters of the National Cancer Registry include also the data on diagnosis and treatment in each reported case. The database can be analysed by a unique software for viewing oncological data (SVOD) developed in the Center of Biostatistics and Analyses of the Masaryk University in Brno. The differences between the results from clinical studies and actual practice as well as variations in treatment results between different hospitals can be envisaged also as equity issue. In the article, there are demonstrated six examples of results achieved in population registry from routine practice which are concordant with conclusions from controlled clinical studies and their meta-analyses. The examples include results of adjuvant chemotherapy in stage II breast cancer and stage III colorectal cancer, preoperative radiotherapy in stage III rectal cancer, adjuvant chemotherapy in stage III malignant melanoma, comparison on various therapies in hepatocellular carcinoma and chemotherapy in stage III-IV soft-tissue sarcomas.

Key words: results from controlled clinical trials, national cancer registry, comparison of survival curves, evidence-based medicine, equity issue

Smyslem organizovaných klinických studií je hledat a potvrdit optimální léčebné postupy za podmínek, které redukuje počet faktorů ovlivňujících výsledky srovnávaných léčebných protokolů. Randomizovaná kontrolovaná klinická studie je pokládána za jediný způsob vědeckého průkazu rozdílnosti výsledků srovnávaných léčebných postupů.

Málo se však hovoří o tom, že jde sice o vědecký nástroj bádání, avšak nikoli jeho konečný cíl. Z chování mnohých výzkumných týmů se zdá, že finálním produktem klinických hodnocení je pouze publikace výsledků kontrolovaných studií v předních odborných časopisech a na velkých mezinárodních konferencích bez zájmu o další implementaci v praxi.

Prezentace výsledků léta běžících a inzerovaných klinických stu-

dií jsou odbornou veřejností s napětím očekávány a mnohdy vedou ke změně doporučených diagnosticko-léčebných postupů či tzv. guidelines, řídí se jimi jednotlivci, celá pracoviště či celonárodní a mezinárodní léčebné protokoly v rutinní praxi. Tento postup je označován za medicínu založenou na důkazech (*evidence based medicine*), v našem případě tedy onkologii založenou na důkazech.

Obvykle nevyčtenou slabinou našeho přesvědčení o vědeckosti tohoto postupu je, že zpravidla končí v polovině cesty. Tedy v přenesení postupu, který se ukázal v kontrolované klinické studii jako výhodnější, do klinické praxe s vírou, že v rutinní praxi bude tento přínos stejný nebo aspoň srovnatelný. Zpravidla chybí zpětná vazba prokazující, jakého přínosu nový postup dosáhl

v rutinní praxi, zatížené vysokou heterogenitou souborů a celou řadou faktorů, které mohou výsledek ovlivnit. Aplikujeme-li tedy výsledky kontrolovaných klinických studií pouze s vírou, že jejich odraz v praxi bude přínosný, nejde nakonec o medicínu založenou za důkazů, nýbrž na důvěře, že podmínky praxe jsou jaksi podobné podmínkám randomizované studie a výsledek musí být podobný. Ve skutečnosti však tímto poskytujeme jen *confidence-based medicine*, nikoli *evidence-based medicine*. Nebo, přesněji řečeno, provozujeme evidence-based medicine v první etapě badání následovanou confidence-based medicine v etapě druhé, tedy aplikační. Můžeme se pak setkat i s tím, že zatímco přibývá úspěšnosti léčby v jednotlivých randomizovaných studiích, nemění se nijak celková úspěšnost léčby v populaci, byť jsou protokoly úspěšné ve studiích respektovány a aplikovány. Zpravidla je nová úspěšnější také nákladnější a po letech tohoto „*evidence-based, but confidence-followed*“ postupu se můžeme dobrat rozporu dramatického nárůstu nákladů léčby (třeba o řád) bez odezvy v pozitivních změnách v populačním registru, tedy v redukci mortality, respektive redukci poměru mezi mortalitou a incidencí.

Ostatně již v úvodu diskusi o významu populačních registrů tvrdíme, že proměny léčebných postupů zpravidla celkovou mortalitu významně neovlivní a k redukci mortality lze přispět pouze časnějším záchytem daného nádoru, tedy populačním downstagingem. Ekonomové a manažeři nezaují medicínskou stránkou hledání lepších léčebných postupů by se pak mohli ptát, proč tedy navrhujeme pro širokou aplikaci léčbu novou a nákladnější, když se neprojeví u většího počtu nemocných v denní praxi. Ptát by se sice mohli, ale zpravidla se neptají. Ekonomika zdravotní péče je totiž, zejména u nás, pojmána pouze deskriptivně jako distribuční či přidělový problém. Analytika v praxi zcela chybí. *Cost/benefit analýzy* jsou u nás extrémně vzácné a dokonce budí dojem něčeho nedůstojného, nehumánního, technokratického a pouze krizového řešení různých propadů či deficitů. Pokud se pak vůbec při hrazení nové a nákladnější léčby argumentuje i něčím jiným než lobistickými tlaky, rétorikou podpory pokroku a modernosti, tak jsou to právě výsledky randomizovaných klinických studií, od nichž se automaticky odvozuje očekávání přínosu v praktických aplikacích.

Jako východisko pro změnu je tento postup asi jediný možný, ale poté již jakýkoli proces validace postupů a verifikace výsledků v praxi zcela chybí. Proces jakoby skončil uvedením protokolu z kontrolované studie do praxe. Validací postupů ovšem míníme ověření, zda v praxi je skutečně uplatňován protokol, použitý v randomizované studii. Teprve v poslední době se začíná věnovat pozornost třeba redukci dávkové intenzity v chemoterapii a zjišťuje se, že jen u části nemocných je dodrženo původní doporučené léčebné schéma jak pokud jde o dávku tak o intervaly podání. Obvykle nejde o svévoli, ale o situace vynucené nežádoucími reakcemi, pestrostí klinických situací, interkurencemi, provozními událostmi nebo i individuální obavou z dalších nákladů při zvládnutí toxických účinků podpůrnou léčbou. Vážne-li již validace nově do praxe přijatého protokolu, pak zcela jistě vážne i verifikace celkových výsledků. Většina pracovišť, snad mimo velká onkologická specializovaná centra, totiž nemá ani vypracovány nástroje pro hodnocení celkových výsledků své léčby a prakticky nepracuje s daty parametrizovanými do té míry, aby tato verifikace byla možná.

Pohybujeme se tak, nejen u nás, ve světě možného, nikoli ideálního hodnocení klinických výsledků. Evidenci čili průkaznost kontrolovaných studií tak následuje konfidence čili důvěra, že to v praxi nakonec také bude nějak přínosné. Tento stav je dán disperzní povahou péče, přehnaným soustředěním odborníků i plátců na dílčí výkon a jen malou koncentrací na celý postup a řešení problému nádorového onemocnění vcelku. Je-li péče disperzní, chaotická a procesně málo provázaná, chybějí také vůle i nástroje k parametrizaci dat, jejich shromažďování a vyhodnocování a tudíž i možnost ověřit dopad nového léčebného protokolu v praxi. Dílčí výsledky relativně malých souborů jsou spíše dílem jednotlivců, ať již entusiastů nebo cílevědomých autorů disertačních, habilitačních či grantových prací. Pokud jejich výsledky potvrdí závěry

kontrolovaných studií, je dopad této podpůrné informace pokládán za nevýznamný, pokud jsou se závěry kontrolovaných studií v rozporu, jsou výsledky zpravidla odmítány jako nereprezentativní a neschopné zpochybnit závěry studie pro praktické aplikace.

Je zřejmé, že praktická studie nemůže nikterak revidovat výsledky randomizované studie. Může však naopak dobře popsat jak se protokol klinické studie chová či uplatní v běžné praxi. Jde ovšem nejen o jeho finální výsledky, nýbrž také o validaci, byl-li vskutku uplatněn beze změn nebo s jakými změnami a v jaké míře. Podmínky praxe nemusejí totiž dovolovat použití stejných kritérií, s nimiž pracoval protokol kontrolované studie v selektované skupině nemocných. Případné změny a odlišnosti pak mohou rozhodovat o konečném souladu nebo diskrepanci mezi výsledky studie a praxí.

Je tedy otázkou, jaký obraz nám dávají jen samotné výsledky klinických studií pro hodnocení posunů v onkologické léčbě v celé populaci, tedy u drtivé většiny nemocných. Vždyť v podmínkách kontrolovaných klinických studií je u nás léčeno jen pár procent nebo spíše promilí nemocných ve vazbě na několik specializovaných pracovišť. Ostatně podobně je tomu ve většině států světa. Přitom léčebné postupy a jejich vývoj jsou diktovány jen výsledky randomizovaných studií, oznamovaných zpravidla na mezinárodních kongresech ASCO či ESMO.

Jakousi paralelní snahou objektivizovat výsledky jednotlivých studií, které mohou být, jak již historie ukázala, i metodologicky chybné či z různých důvodů zkresené, jsou *metaanalýzy*. Srovnáváním výsledků více randomizovaných klinických studií v kompilované podobě, která zvažuje i odlišnosti souborů či metodiky, se lze dopracovat k jakémusi vyššímu stupni evidence. Zpravidla však řešení určitého problému končí tak, že část studií nový postup potvrdila a část nikoli. Statistický přínos je pokládán za tím významnější, čím větší jsou komparované soubory. Na druhé straně je zřejmé, že přínos je nepochybný právě tam, kde rozdíl je markantní již v souborech malých. Mnohé studie byly ukončeny právě proto, že výsledky v jednom rameni protokolů byly natolik lepší či horší, že další randomizace by jednu skupinu nemocných zcela jasně znevýhodňovala a pokračovat ve studii by bylo neetické.

Ani metaanalýza ovšem nikterak neobjasňuje přínos nových postupů v praxi, tedy v neselektované populaci. Vyznívají-li její závěry jednoznačně, zvyšuje snad šanci, že přínosu bude dosaženo i v populaci jako celku. O míře tohoto přínosu však informaci nepodává. Neselektovanou populaci samozřejmě není miněno bezbřehé užívání protokolu studie, nýbrž respektování indikačních kritérií vyplývajících ze zadání kontrolovaného klinického hodnocení.

Mnohými je však dosud stále pokládáno spíše za troufalé hledat dopady studiem doporučených léčebných postupů v populačním registru. Ostatně i v mnoha velmi rozvinutých zemích tato možnost ve skutečnosti ani není. Populační registry těchto zemí jsou nezfídka limitované geograficky i obsahově. Nejčastěji sledují jen základní epidemiologické údaje bez informací o stadiu onemocnění a provedené léčbě, jsou vázány jen na některé oblasti v blízkosti vlivu velkých onkocenter nebo dokonce zcela chybí. Proto také chybí analýza faktorů, které mohou výsledky dosahované v populačních registrech ovlivňovat, samozřejmě za předpokladu, že jsou srovnávány soubory stejné věkové distribuce a onemocnění stejného stádia pokročilosti. Na druhé straně umožňují populační registry, zahrnují-li údaje o provedené léčbě, analyzovat řádově větší soubory než klinické studie a vypovídat o jejich praktickém medicínském i ekonomickém dopadu. Mohou také vypovídat o regionálních rozdílech v dostupnosti a provádění onkologické léčby a dotýkají se tak citlivého *aspektu ekvity* čili dostupnosti péče srovnatelné kvality a výsledku. Vždyť za předpokladu stejných plateb, ať již cestou pojištění, přímých nebo jejich kombinací, má mít pacient jako zákazník právo na výsledek srovnatelný s jinými stejně postiženými a stejně placenými. A jsou-li nalezeny markantní rozdíly mezi regiony či pracovišti, pak je třeba hledat a napravovat jejich příčiny.

Problém ekvity výsledků léčby, ostatně i srovnatelných rizik a nákladovosti, bude v informační společnosti narůstat tak jak se zpřístupňují informace o všem, včetně medicíny. Translace

výsledků studií získaných za nadstandardních podmínek a variabilita výsledků postupů aplikovaných ve standardních podmínkách praxe jsou jedním z aspektů pootevření Pandoriny skříňky ekvity medicínské praxe. Není zvláštní ochota ji otevřít a tudíž ani zvlášť dobrá metodická připravenost. Je však nepochybné, že postupně otvírána bude nejen z důvodů morálně-etických, nýbrž i ryze vědeckých a samozřejmě také ekonomických. Které důvody budou hlavní si netroufám předjímat, soudím však, že stejně jako v jiném lidském konání to budou nejprve důvody ekonomické, byť by se chtělo chlubit důvody jinými.

V zásadě je potěšující, že v České republice máme k analýzám populačního onkologického registru přístup nejen po stránce hrubé epidemiologie. Zasloužily se o to vědomě i nevědomě generace lékařů z praxe a pracovníků okresních registrů, které po tři desetiletí pod tlakem zákonem dané povinnosti hlášení zhoubného novotvaru v horší či lepší podobě, za lepších či horších organizačních podmínek, shromažďovaly informace o diagnostice a léčbě nádorů, které jsou nyní k dispozici v parametrizované databázi Národního onkologického registru. Způsob parametrizace ani její interpretace a dodržování nejsou dosud zdaleka ideální a zaslouhují další kultivaci. Přesto však umožňuje hrubé pohledy na dopad aplikací velkých mezinárodních klinických studií v české populaci onkologicky nemocných. Jde zejména o přínos doporučených adjuvantních postupů nebo naopak jejich zpochybnění. Umožňuje také sledovat posuny ve výsledcích léčby v čase a případně i benchmarking vlastního pracoviště vůči pracovištím jiným. Ze společenských důvodů je však v praxi dosud možné spíše jen porovnávání vlastních výsledků vůči průměru republiky a kraje, což nevyvolává tolik vášni jako srovnávání a rozdělování pracovišť podle výsledků péče.

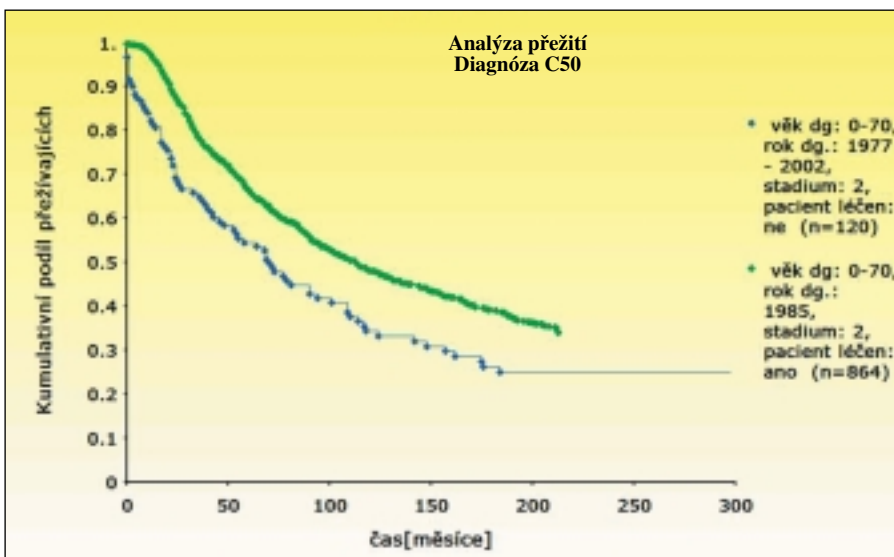
Hovoříme-li však o odrazu klinických studií v populačním registru, nemusíme řešit problém benchmarkingu a ekvity, nýbrž pouze soulad či diskrepanci výsledků s doporučenou taktikou léčby v celostátním kontextu. Přestože databáze NOR by byla i připravena pro jemnější srovnávání léčebných schémat, nemají dodávané údaje takovou rozlišovací schopnost, aby bylo možno vyhodnotit třeba jednotlivé režimy adjuvantní chemoterapie. Další kultivace této parameterizace záznamů o diagnostice a léčbě bude již nutná na jiné, modernější organizační bázi při menším počtu více specializovaných a za své výsledky zodpovědných pracovišť komplexní péče - onkocenter nebo onkologických kooperativních skupin. Tak je tomu již v dětské onkologii nebo v hematologii. Tudíž jsou tam mnohem podrobněji pracovány a vyhodnocovány specifické záznamy pro léčbu leukemií, lymfomů nebo mnohočetného myelomu nebo protokoly mezinárodně přijaté v dětské onkologii. Jde ovšem o oblast jen 5-7% onkologických diagnóz. „Zbytkový“ prostor oněch 95% solidních zhoub-

ných nádorů je zapotřebí ještě kultivovat a sjednocovat ideově i organizačně.

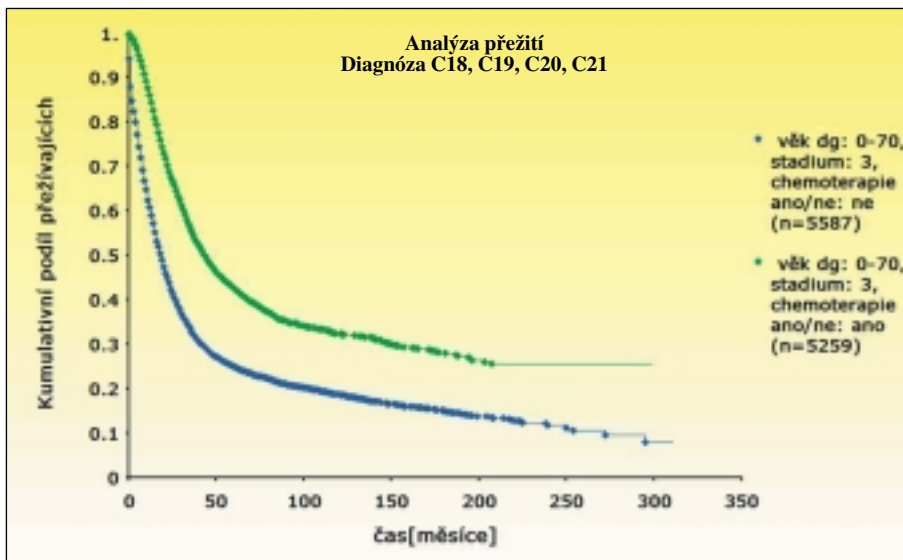
Při zohlednění všech uvedených relativizujících kritérií a překážek lze pokládat za velký úspěch, že se podařilo nalézt v populačním registru NOR potvrzení některých nejobecnějších doporučení vzešlých z mnohaletých a opakovaných klinických studií, které se týkají především adjuvantní léčby v návaznosti na chirurgický výkon. Toto pozitivní zjištění je však nuto vidět také z kvantitativního pohledu a v závislosti na investovaných prostředcích.

Kritický pohled bychom získali, kdybychom celkový přínos poměřovali rozdíly ve výsledcích péče mezi jednotlivými pracovišti. Například činí-li přínos adjuvantního postupu v populačním registru něco kolem 10%, bývají rozdíly ve shodném postupu u stejného stadia a při stejném věkové struktuře nemocných mezi jednotlivými pracovišti 20-40%. Je zřejmé, že zásadní roli hraje počet případů daného onemocnění, řešený ročně, týdně či měsíčně, a při poklesu pod jakousi kritickou mez, obvykle pod 50-100 případů ročně, se výsledky péče markantně zhoršují. To je hlavní důvod pro snahu o integraci léčebných postupů do méně míst, do vybavených, kvalifikovaných a na problematiku soustředěných center. Ostatně o prospěšnosti soustředování transplantací orgánů, kardiokirurgie či neurochirurgie již nikdo nepochybuje. Byla také deklarována koncepce traumacenter. Pouze při menším množství pracovišť, zabývajících se většími počty případů daného onemocnění (*high volume centers*), lze dále kultivovat parameterizaci medicínských záznamů, sběr i vyhodnocování dat a provádět komparace s výsledky kontrolovaných studií jako doklad jejich dopadu v široké praxi.

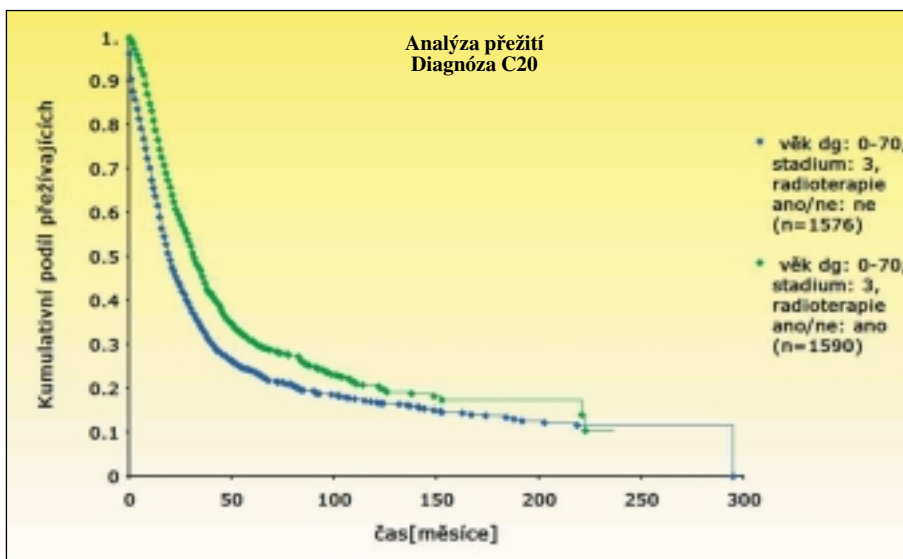
Na závěr tohoto pojednání uvádím šest pozitivních případů koncordance doporučení z kontrolovaných klinických studií a jejich metaanalýz s výsledky dosahovanými v české populaci onkologicky nemocných jak je nabízí Národní onkologický registr (NOR) České republiky a jeho zpracování Systémem vizualizace onkologických dat (SVOD) v originálním softwarovém řešení Centra biostatistiky a analýz Masarykovy univerzity v Brně. Ukazují tím, že metodicky je proveditelné hledat dopad klinických studií v populačních onkologických registrech, zahrnují-li potřebné parametry a mají-li k dispozici podobné analytické nástroje jakým je SVOD. Případy diskordance výsledků klinických studií a praktických aplikací zde neuvádím záměrně, protože ty by vyžadovaly podrobnější analýzy a komentáře důvodů nesouladu, přesahující rámec tohoto příkladového sdělení. Ačkoli v NOR jsou k dispozici údaje od roku 1977, uvádím zde hodnocení souborů z let 1995-2002, které již vesměs odrážejí současný stav léčby po akceptaci výsledků mezinárodních kontrolovaných studií.



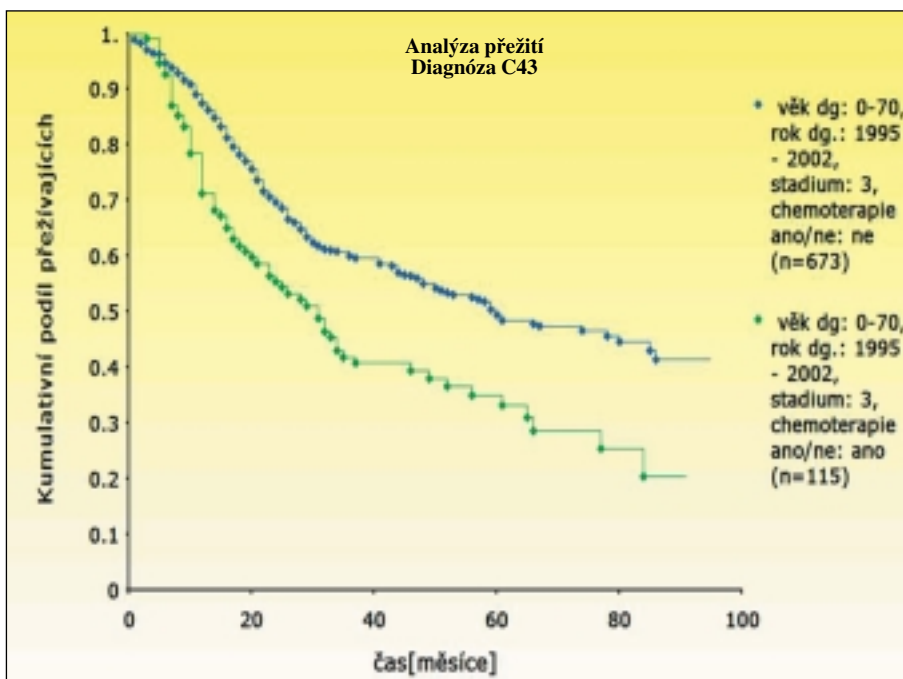
Obr. 1: Přínos adjuvantní chemoterapie ve II. klinickém stadiu karcinomu prsu byl opakovaně po léta prokazován v kontrolovaných klinických studiích EBCTCG (Early Breast Cancer Trialists' Cooperative Group) i v nejnovějších metaanalýzách v zahraničí (1) a nalézáme ho v přesvědčivé podobě také v analýze údajů NOR.



Obr. 2: Přínos adjuvantní chemoterapie kolorektálního karcinomu ve stadiu III (Dukes C), vyplývající z posledních přehledů klinických studií (2, 3). lze demonstrovat také v našem populačním onkologickém registru.

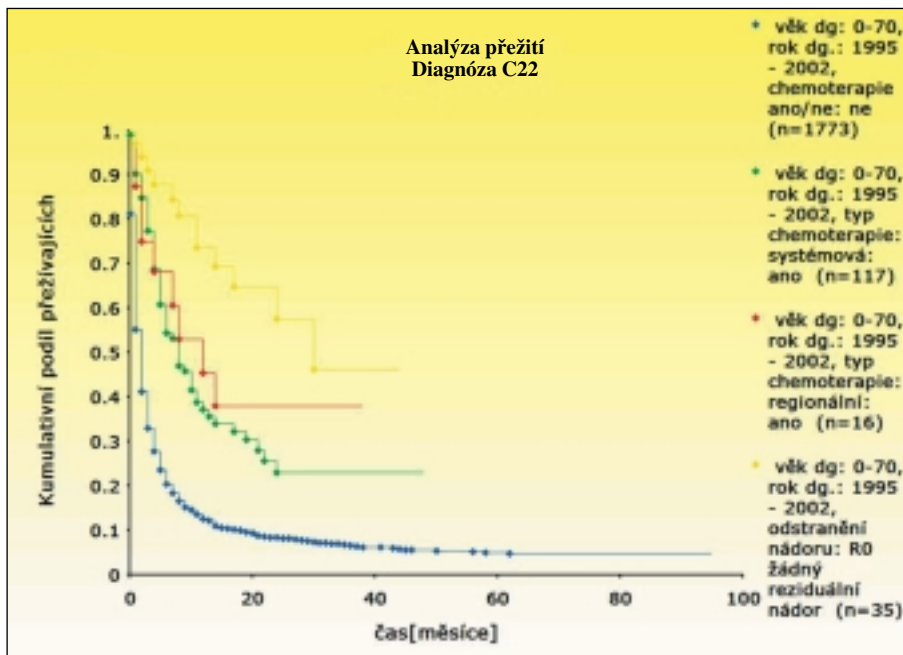


Obr. 3: Poslední souhrnné přehledy klinických studií (3, 4) opětovně potvrzují přínos předoperační radioterapie karcinomu rekta stadia III (Dukes C) a toto zlepšení léčebných výsledků předoperační radioterapií lze demonstrovat také v populační databázi NOR.



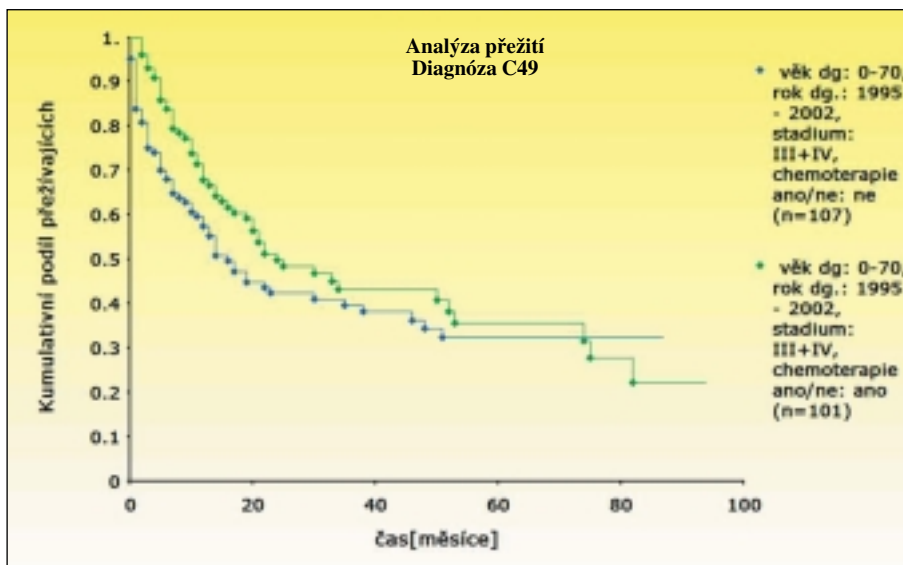
Obr. 4: Adjuvantní chemoterapie u maligního melanomu s uzlinovým postižením, tedy klinického stadia III, svůj přínos v klinických studiích neprokázala. Dokonce ani adjuvantní podání interferonu není dosud jednoznačně doporučeno v rutinní praxi mimo podmínky kontrolovaných klinických hodnocení (5). Podobnou absenci přínosu nacházíme také v analýze údajů NOR. Je naopak zneklidňující, že zhoršené výsledky vykazuje v populačním registru ve srovnání s chemoterapií neléčenými skupina nemocných chemoterapií léčená. Může jít o důsledek selekce pokročilejších případů masivnějšího uzlinového postižení ve stadiu III pro chemoterapii. Mohlo by však jít také o nežádoucí důsledky perturbujícího vlivu protinádorově neúčinného podávání cytostatik na chemoresistentní nádory ve smyslu akcelerace jejich růstu, což je problematika, které se současná onkologie poněkud vyhýbá. Nálezy v populačním registru tak mohou zpětně přinášet nové náměty i pro experimentální studie.

Analyza přežití Diagnóza C22



Obř. 5: Možnosti léčby hepatocelulárního karcinomu závisí na rozsahu postižených jater i zkušenostech pracoviště. Ve spektru léčebných možností se jako neúčinnější nabízí chirurgická resekce, je-li ovšem možná, a dále se diskutuje, zda dát přednost regionální nebo systémové chemoterapii (6). Z rozboru případů v populačním registru NOR vyplývá, že drtivá většina případů (91%) hepatocelulárního karcinomu není v praxi chemoterapii léčena vůbec. Resekci nádoru jater s mikroskopicky negativními okraji resekátu (R0) podstoupil jen 2% nemocných, systémovou chemoterapii 6% a regionální chemoterapii méně než 1% nemocných. Léčebné výsledky podle očekávání ukazují, že kompletní chirurgická resekce přináší výsledky nejlepší. Regionální chemoterapie má mírně lepší výsledky než chemoterapie systémová, což může souviset i selekci vhodnějších a méně pokročilých případů. V případech neléčených jsou však výsledky mnohonásobně horší a pětileté přežití činí jen pár procent. Spíše než srovnání výsledků léčebných metod zde analýza registru ilustruje pravý stav léčebné péče o nemocné s hepatocelulárním karcinomem v populaci. Samozřejmě, že součástí detailní analýzy by mohlo být také srovnání věkového rozložení léčených a neléčených případů, srovnání stavu v regionech a vliv pracoviště na proporce aktivní léčby.

Analyza přežití Diagnóza C49



Obř. 6: Výsledky léčby sarkomů měkkých tkání jako nádorů méně častých se mohou lišit v závislosti pracoviště a racionálně řízené péči (7), užívané klasifikaci léčebné odpovědi (8) i aktuální chemosenzitivitě nádoru (9). Většina kontrolovaných studií deklaruje mírný přínos chemoterapie ve stádiích III-IV. V populačním registru bylo chemoterapii léčeno v těchto stádiích polovina nemocných se sarkomy, avšak přínos této léčby činí méně než 10% do 2 let a po 5.roce se již zcela vytrácí. Nepochybně je potřeba odlišné a nové, lépe diferencující strategie pro indikace i organizaci léčby nemocných se sarkomy (10).

Literatura:

1. Van Nes JG, Nortier WR, van de Velde CJ.: The 4th large meta-analysis of all trials of the treatment of operable breast cancer: increased survival after a longer follow-up, Ned Tijdschr Geneesk. 2005 Sep 3; 149 (36): 1978-80.
2. Bretta GD, Milesi L, Pessi MA, Mosconi S, Labianca R.: Adjuvant treatment of colorectal cancer. Surg Oncol. 2004 Aug-Nov;13(2-3):63-73.
3. Hobday TJ.: An overview of approaches to adjuvant therapy for colorectal cancer in the United States. Clin Colorectal Cancer. 2005 Apr; 5 Suppl 1: S11-81
4. Baxter NN, Moris AM, Rothenberger DA, Tepper JE.: Impact of pre-operative radiation for rectal cancer on subsequent lymph node evaluation: a population-based analysis. Int J Radiat Oncol Biol Phys. 2005 Feb 1; 61(2):426-31.
5. Lawson DH.: Choices in adjuvant therapy of melanoma. Cancer Control. 2005 Oct;12(4):236-41.
6. Tang ZY: Hepatocellular carcinoma : cause, treatment and metastasis. World J Gastroenterol. 2001 Aug;7(4):445-54.
7. Gross CP Krumholz HM.: Impact of managed care on cancer trial enrollment. J Clin Oncol. 2005 Jun 1;23(16):3811-8.
8. Therasse P, Le Cesne A, Van Glabbeke M, Verweij J., Judson I; for the EORTC Soft Tissue and Bone Sarcoma Group: RECIST vs. WHO: prospective comparison of response criteria in an EORTC phase II clinical trial investigating ET-743 in advanced soft tissue sarcoma. Eur J Cancer. 2005 Jul;41(10):1426-30.
9. Lenhardt M., Muehlberger T, Kuhnen C, Brett D, Steinhaus HU, Jafari HJ, Steinstrasser L., Muller O, Homann HH.: Feasibility of chemosensitivity testing in soft tissue sarcomas. World J Surg Oncol. 2005 Apr 18;3(1):20.
10. Hartmann JT., Patel S.:Recent developments in salvage chemotherapy for patients with metastatic soft tissue sarcoma. Drugs. 2005;65(2): 167- 78.