proLékaře.cz | 3.7.2024



# Analýza dat v neurologii

## XLVIII. „Rare disease assumption“ v asociačních studiích

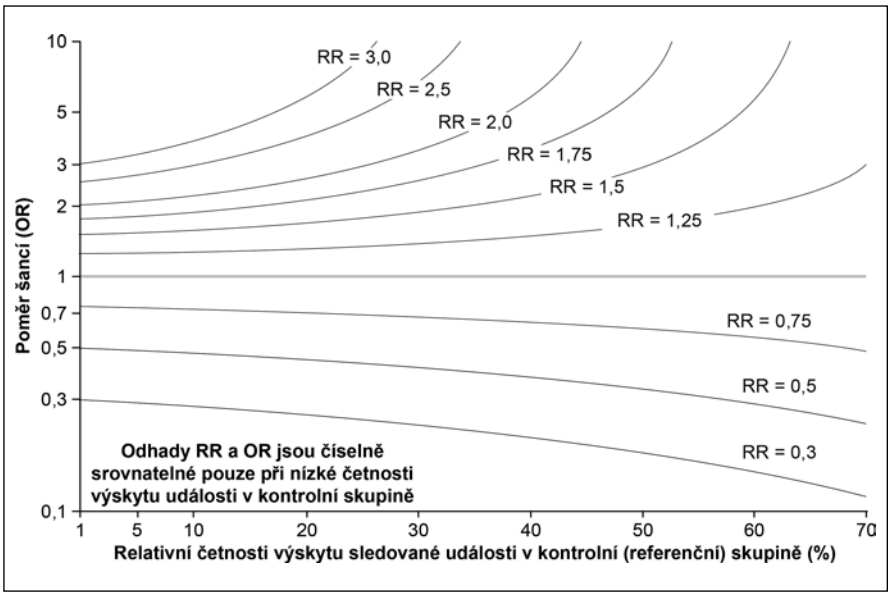
L. Dušek, T. Pavlík,
J. Jarkovský, J. Koptíková
Institut biostatistiky a analýz
Masarykova univerzita, Brno

✉
doc. RNDr. Ladislav Dušek, Dr.
Institut biostatistiky a analýz
MU, Brno
e-mail: dusek@iba.muni.cz

Již delší dobu se v tomto seriálu věnujeme různým aspektům odhadu poměru šancí a relativního rizika jako statistik kvantifikujících vztah „expozice-následek“ v asociačních studiích. V minulých dílech jsme rozebrali různé experimentální plány těchto studií z hlediska jejich pozitiv a negativ, přičemž u všech typů experimentálních plánů byl jako zvláště významný faktor vyzdvižen způsob náběru vzorku. Ten rozhoduje o tom, do jaké míry jsou data vstupující do studie populačně reprezentativní a do jaké míry je bude možné využít pro relevantní odhady incidence a prevalence nemoci. Pro čtenáře byl možná překvapující fakt, že vlastní odhad těchto epidemiologických parametrů většinou není prvořadým cílem asociačních studií. Tyto studie primárně posuzují míru vztahu mezi expozicí a následkem a snaží se ji co nejsprávněji kauzálně interpretovat. Tohoto cíle můžeme dosáhnout například již cíleným náběrem poměrně malých kohort kontrol a pacientů v nemocnicích s tím, že takové vzorky ale většinou nemohou vypovídat o reálném výskytu dané nemoci v populaci. A tedy, ačkoli zcela reprezentativní a náhodně nabraný vzorek nadále zůstává nejvíce ceněným standardem této oblasti medicíny i statistiky, realita mnoha publikovaných observačních studií je jiná. Tento fakt má významné důsledky pro vzájemnou srovnatelnost či vzájemné extrapolace různých odhadů relativního rizika (*RR*) nebo poměru šancí (*OR*). Problému srovnatelnosti odhadů *RR* a *OR* budeme krátce věnovat tento díl seriálu.

Problém vzájemné srovnatelnosti odhadů *OR* a *RR* do značné míry odpadá u retrospektivních studií (typicky studie případů a kontrol, *„case-control“*), u kterých nemá výpočet *RR* interpretační význam a může být zavádějící (viz např. díl XLVI). Tyto experimentální plány vedou k hodnocení vztahu expozice-následek pomocí poměru šancí, jehož hodnoty jsou srovnatelné s relativním rizikem pouze při studiu velmi vzácných jevů (viz díl XLVI, příklady 1 a 2). Nicméně tento jev zaslouženě poutá pozornost, neboť aproximace odhadu relativního rizika z hodnot poměru šancí může být u studie případů a kontrol velkou výhodou; např. v situaci, kdy potřebujeme srovnat závěry takové studie s prospektivními kohortovými studiemi, kde je hlavním výstupem odhad *RR*. Společným prvkem srovnatelnosti *RR* a *OR* je tedy prevalence sledovaného jevu (nemoci) v cílové populaci. Hodnota prevalence určuje velikost vzorku, který bude třeba k záchytu dostatečného počtu jedinců s daným jevem. Je třeba si uvědomit, že s klesající prevalencí jevu sice logicky rostou požadavky na velikost vzorku, ale pouze do chvíle, kdy se stane nereálným nabírat do asociační studie jedince „čistě náhodně z populace“. Experiment je v takové situaci nucen pracovat s jinak definovanými skupinami, kde je výskyt jevu častější, např. pacienti se vzácnou nemocí nabíraní přímo ze specializovaných klinických center a nikoli z běžné populace.

Hodnota prevalence jevu v kontrolní skupině kromě náběru zásadně ovlivňuje i hodnotu odhadovaného relativního rizika. V příkladech dílu XLVI jsme doložili, že s rostoucí prevalencí znaku se hodnota *RR* více odlišuje od hodnoty *OR* a klesá. Na obr. 1 zde doplňujeme přehledný nomogram publikovaný Zhangem a spoluautory v roce 1998, který potvrzuje významný vliv prevalence znaku v kontrolní skupině na vztah mezi hodnotami *OR* a *RR*. Prevalence znaku v populaci je na obrázku vyjádřena jako procento výskytu osob s daným znakem v kontrolní skupině (ramenu) asociační studie. Z obr. 1 vyplývají dva velmi významné závěry:



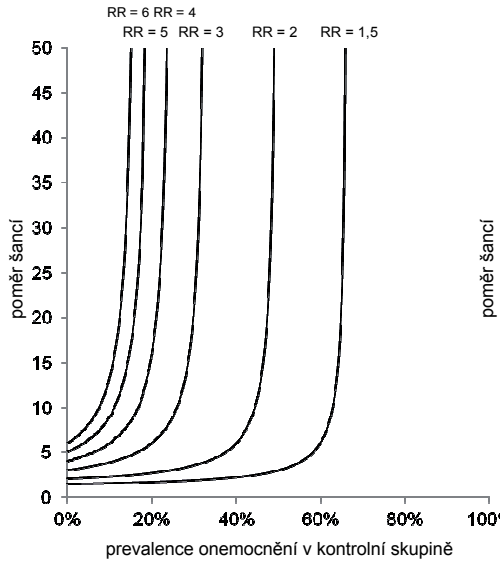

Obr. 1. Srovnání hodnot *RR* a *OR* při různé četnosti výskytu události v kontrolní skupině prospektivních nebo průřezových asociačních studií.

proLékaře.cz | 3.7.2024

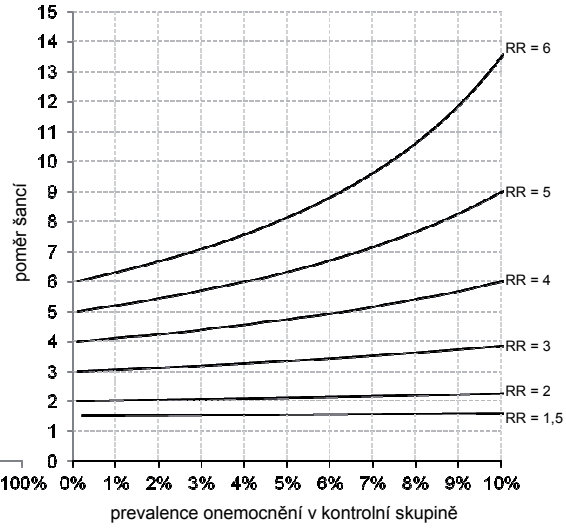


Tzv. předpoklad vzácného onemocnění (*rare disease assumption*) je matematickým předpokladem využívaným v hodnocení asociačních studií případů a kontrol. Jde o předpoklad, že při nízké prevalenci onemocnění lze relativní riziko aproximovat pomocí poměru šancí, který se u tohoto typu studií používá pro hodnocení vztahu mezi expozicí a výskytem onemocnění. Tento postup se jeví jako výhodný mimo jiné i proto, že studie případů a kontrol jsou relativně laciné a časově méně náročné než kohortové prospektivní studie, které umožňují počítat vztah mezi expozicí a výskytem onemocnění pomocí relativních rizik. Vyjádření vztahu expozice a onemocnění pomocí relativního rizika je vhodnější k interpretaci. Řada autorů považuje prevalenci 10 % a méně za dostatečně vzácný výskyt onemocnění, aby tento předpoklad byl naplněn. Jak si však ukážeme v příkladu, prevalence onemocnění není jediným faktorem ovlivňujícím vztah mezi relativním rizikem a poměrem šancí.

Graf ukazuje vztah mezi poměrem šancí a relativním rizikem v závislosti na prevalenci onemocnění v kontrolní skupině. Při vysoké prevalenci je patrný významný rozdíl mezi poměrem šancí a relativním rizikem.




Při omezení grafu na prevalenci do 10 % je patrné, že vztah mezi relativním rizikem a poměrem šancí závisí i na absolutní hodnotě relativního rizika; při prevalenci 10 % a při *RR* = 1,5 a *RR* = 2 jsou rozdíly oproti *OR* zanedbatelné, při *RR* = 6 ale již *OR* dosahuje hodnoty kolem 13,5; a *OR* jako odhad *RR* tak nadhodnocuje výsledné relativní riziko.




**Shrnutí:** Předpoklad vzácného onemocnění umožňuje odhad relativního rizika aproximovat z odhadu poměru šancí získaného ve studii případů a kontrol. To může být výhodné například při metaanalýze různých publikovaných studií, pokud srovnáváme výstupy kohortových studií a studií případů a kontrol.

Problémem uplatnění tohoto předpokladu je však fakt, že neexistuje exaktně daná hranice pro nezkreslenou aproximaci hodnot *RR* z odhadu *OR* a kromě vlastní prevalence onemocnění je vztah mezi *RR* a *OR* ovlivněn i absolutní hodnotou relativního rizika a poměru šancí v dané studii. **Předpoklad vzácného onemocnění je tak vhodné aplikovat pouze v případě, kdy je ověřena jeho platnost v konkrétní hodnocené situaci.**

Příklad 1. Předpoklad vzácného onemocnění při odhadu relativního rizika *(RR)* a poměru šancí *(OR)* v asociačních studiích.

1. Odhady *RR* nebo *OR* dosažené ve studiích lišících se významně prevalencí znaku v kontrolní populaci jsou vzájemně obtížně srovnatelné a jejich přímočaré spojování (např. za účelem metaanalýz) není přípustné.
2. Hodnoty *RR* se blíží hodnotám *OR* pouze při vzácném výskytu sledovaného znaku v populaci.

Především druhý z uvedených závěrů má podstatný dopad na srovnatelnost hodnot *OR* a *RR* pocházejících z různých studií, lišících se nejen prevalencí jevu v kontrolní skupině, ale i experimentálním plánem. Srovnatelnost a případně i číselný přepočet hodnot *OR* a *RR* jsou možné pouze při vzácném výskytu sledovaného jevu; tento předpoklad či podmínku literatura nazývá *„rare disease assumption"*.

Odbočme nyní ve výkladu a krátce se zastavme u pojmu vzácná onemocnění. Jde o epidemiologický termín označující choroby s prevalencí nižší než 5 pacientů na $10^4$ osob v populaci (někdy též vyjadřováno jako nižší než 1 nemocný na 2 000 osob, tedy 1 : 2 000). Přepočteme-li uvedenou hodnotu na českou populaci, pak jde o choroby s absolutním počtem pacientů v prevalenci menším než 5 000 osob. Takto kvantitativně definovaná hranice je používána zejména v Evropě, v jiných kontinentech mohou být definiční hodnoty mírně odlišné. Rovněž samotný výskyt nemoci se může v různých populacích významně lišit a onemocnění, které je vzácné v našich podmínkách, může být na jiných kontinentech velmi četné. Čistě epidemiologická definice vzácného onemocnění může být doplněna i klinickými hledisky, např. v EU se mezi vzácná počítají zejména onemocnění, která pacienty ohrožují na životě nebo je chronicky poškozují. Více informací lze nalézt například na portálu http://ec.europa.eu/health/rare_diseases.

Vzácná onemocnění bývají poměrně často předmětem nejrůznějších studií, neboť paradoxně v celkovém souhrnu vzácná nejsou. Jen pro země EU dostupné odhady ukazují, že nějakou nemocí z této kategorie trpí přibližně 30 mil. obyvatel, což tvoří až 8 % populace. Vzácná onemocnění jsou tedy velmi různorodou skupinou několika tisíc nemocí, které spojuje poměrně přesně epidemiologicky kvantifikovaná míra výskytu. Zde je ale nutné zdůraznit, že toto vymezení se asociačních studií a možnosti srovnat různé odhady *OR* a *RR* týká pouze částečně. Hodnoty *OR* a *RR* se vzájemně přiblíží při vzácném výskytu jevu (obr. 1), ale tím je zejména myšlen vzácný výskyt jevu v kontrolní skupině konkrétní získané asociační tabulky četností 2 × 2. Jde tedy spíše o výsledek konkrétních vzorkovacích plánů a konkrétního výpočtu než o naplnění definice vzácného onemocnění z epidemiologického hlediska. Proto také hovoříme o naplnění **předpokladu vzácného onemocnění**.

Jak již bylo řečeno, tento předpoklad a aproximace hodnot *RR* z odhadů *OR*

proLékaře.cz | 3.7.2024



jsou největším přínosem u studií případů a kontrol zaměřených na onemocnění či jevy, které jsou v cílové populaci vzácné. V takové situaci se odhad poměru šancí blíží odhadu relativního rizika a oba odhady lze tedy považovat za vzájemně zaměnitelné. V literatuře je za hranici, kdy můžeme uplatnit předpoklad vzácného onemocnění, nejčastěji (ale ne vždy) považován 10% a nižší výskyt zkoumaného jevu v kontrolní populaci. Numerická blízkost odhadů *OR* a *RR* je v této oblasti patrná i z obr. 1.

Předpoklad vzácného onemocnění má nezanedbatelný význam při metaanalýzách většího počtu heterogenních studií, neboť je-li uplatnitelný, lze vzájemně srovnávat i poměrně různorodé publikace prezentující odhady *OR* či *RR*. A to i v případě, kdy vedle sebe stojí výstupy retrospektivní studie případů a kontrol a výstupy prospektivní studie kohortové. Abychom ale byli spravedliví, musíme zdůraznit, že tento postup má v literatuře i své hlasité kritiky, kteří poukazují zejména na to, že není stanovena exaktní hranice prevalence, při které lze předpoklad vzácného onemocnění bezpečně uplatnit. Platí sice obecné pravidlo, že čím je prevalence nemoci nižší, tím více se *OR* blíží k hodnotám *RR*, hranice 10% prevalence však bývá kritizována jako příliš vysoká. Abychom se vyhnuli zkreslení výsledků a zavádějící interpretaci, je vhodné vzájemnou zaměnitelnost *OR* a *RR* v praxi omezit na 5% a nižší výskyt zkoumaného jevu v kontrolní populaci. Pro zájemce zde nabízíme v seznamu literatury několik prací, kde je tato problematika probírána podrobněji. Dále tomuto problému věnujeme příklad 1, který dokládá, že kromě vlastní prevalence sledovaného jevu závisí vzájemná relace odhadů *OR* a *RR* také na absolutní hodnotě těchto statistik u konkrétní studie. Při vyšších hodnotách *RR* či *OR* roste pravděpodobnost zkreslení, pokud z odhadu *OR* aproximujeme hodnotu *RR*. Proto je doporučováno nepoužívat předpoklad vzácného onemocnění slepě jako všeobecně platné pravidlo, ale každou aplikaci řádně zvážit i z hlediska možných zkreslení a chyb. Případně platnost pravidla ověřit. Problémům ověření vzájemné extrapolace odhadů *OR* a *RR* se budeme rovněž dále věnovat v příkladech dalšího dílu našeho seriálu.

## Literatura

Greenland S, Thomas DC. On the need for *R* the rare disease assumption in case-control studies. Am J Epidemiol 1982; 116(3): 547–553.

Greenland S, Thomas DC, Morgenstern H. The rare-disease assumption revisited. A critique of „estimators of relative risk for *R* case-control studies". Am J Epidemiol 1986; 124(6): 869–883.

Zhang J, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. JAMA 1998; 280: 1690–1691.

Cummings P, Koepsell TD. On the need for *R* the rare disease assumption in some case-control studies. Inj Prev 2001; 7(3): 254.