proLékaře.cz | 16.8.2025

Analýza dat v neurologii

L. Vybrané komentáře k odhadům a interpretaci poměru šancí

L. Dušek, T. Pavlík,
J. Jarkovský, J. Koptíková

Institut biostatistiky a analýz
Masarykova univerzita, Brno

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

Tento jubilejní díl seriálu budeme věnovat dílčím komentářům zaměřeným na vybrané problémy spojené s odhady šance (*O*) a poměru šancí (*OR*) v asociačních studiích. Uzavíráme tak kapitoly, na které nebyl dostatečný prostor v předchozích dílech seriálu, a také odpovídáme na nejčastější otázky, které v této souvislosti zaznívají v literatuře či na přednáškách.

Pro srozumitelnost v úvodu opakujeme nejjednodušší tabulku četností 2 × 2 s typickým značením četností v jednotlivých polích *a, b, c, d* (viz též díl seriálu 49) (tab. 1).

Jak postupovat v případě, kdy se v tabulce četností objeví políčko s nulou?

Pro výpočet poměru šancí je tato situace vážnou komplikací, což je patrné již z výše uvedených vztahů pod tabulkou četností. Pokud nulová četnost nastane v buňkách *b, c,* pak poměr šancí nelze spočítat, neboť nemůžeme dělit nulou. Jsou-li nulové buňky *a, d*, pak *OR* sice spočítat můžeme, ale nevyhnutelně vyjde nulové, což nedává interpretační smysl. Do publikací v obou případech píšeme poznámku „*OR* not defined due to zero cells in frequency table". Nastane-li taková situace, je třeba prověřit, zda skutečně jde o kategorie vzácné až nemožné, kde je nulová četnost objektivní fakt, anebo jde o chybu vzorkování. S takovou chybou se setkáváme především u malých vzorků, které snadno podhodnotí výskyt některé vzácnější kategorie.

Jak se u asociačních prospektivních a retrospektivních studií liší cílový parametr, end-point?

V zásadě se neliší, oba experimentální plány sledují stejný cíl. Jednoduše řečeno, jde o to posoudit vliv expozice (tedy experimentálního zásahu, léčby, vliv nějakého znaku, prediktoru, rizikového faktoru) na výskyt zkoumaného jevu (typicky nemoci). Výskyt nemoci je tedy cílovým parametrem, přičemž ho měříme jako relativní četnost ve skupině exponovaných a neexponovaných jedinců. Rozdíl mezi prospektivním a retrospektivním plánem je pouze ten, že se na daný problém dívají z opačné strany. Pouze prospektivní plán ovšem umožňuje reprezentativní odhad prevalence zkoumaného jevu, proto je možné u něj odhadovat relativní riziko (*RR*) jeho výskytu. Retrospektivní plány již při zařazování jedinců do studie rozhodují o poměru případů a kontrol, relativní výskyt případů tedy není nezávislý odhad, ale výsledek náběru probandů.

Tab. 1. Výpočet relativního rizika (*RR*) a poměru šancí (*OR*) na základě tabulky četností 2 × 2.

Dělení na případy a kontroly dle výskytu sledovaného jevu (nemoci; u retrospektivní studie případů a kontrol jde o první kritérium výběru osob)	Dělení dle expozice (vlivu sledované události) jedinců v minulosti (u retrospektivní studie zjištěno zpětně)	
	expozice (událost): ano	expozice (událost): ne
případy	*a*	*b*
kontroly	*c*	*d*

- Podíl exponovaných osob u případů a kontrol: $a/(a + b)$; $c/(c + d)$.
- Šance výskytu jevu u exponované a neexponované skupiny: $O_E = a/c$; $O_{NE} = b/d$.
- Výpočet *OR* provádíme podle vztahu: $OR = a*d/b*c$.
- *RR* nelze u retrospektivní studie přímo odhadnout (výběr a priori definuje poměr případů a kontrol, jejich vzájemný poměr tedy není reprezentativním odhadem).
- Pokud by šlo o studii prospektivní, pak bychom nejprve provedli expozici (nebo vybrali jedince s událostí a bez) a sledovali v čase výskyt daného jevu (nemoci);
  - riziko výskytu sledovaného jevu u exponované a neexponované skupiny bychom v takové situaci kalkulovali: $R_E = a/(a + c)$; $R_{NE} = b/(b + d)$.
  - Výpočet *RR* bychom potom provedli podle vztahu: $RR = [a/(a + c)]/[b/(b + d)]$.

Existuje nějaké jednoduché pravidlo, kdy použít *OR* a kdy *RR*?

Ano, existuje. Jak bylo vysvětleno v dílech 48 a 49 našeho seriálu, nejlepší doporučení jsou tato:

- *RR* nemá smysl a může být silně zavádějící u retrospektivních studií (studie případů a kontrol, retrospektivní kohortové studie); zde je naopak dobře interpretovatelný odhad *OR*.
- U prospektivních observačních studií (kohortové, průřezové) lze sice odhadnout *OR* i *RR*, z hlediska interpretace je ale lepší *RR*, neboť pracuje s reprezentativním odhadem rizika (pravděpodobnosti) výskytu jevu v exponované a neexponované skupině; to samé platí u prospektivních randomizovaných studií.
- Rovněž následné metaanalýzy, je-li to možné, srovnávají preferenčně odhady *RR*. U studií, kde to jde, je možné hodnotu *RR* aproximovat (přepočítat) z odhadu *OR* (např. při splnění předpokladu vzácných onemocnění).

proLékaře.cz | 16.8.2025

Vstupní tabulka dat shrnuje výsledky studie zkoumající vliv určité léčby (expozice) na dosažení léčebné odpovědi (úspěch/neúspěch). Předpokládejme prospektivní pozorovací studii, u které má smysl současně hodnotit poměr šancí (*OR*) i relativní riziko (*RR*).

Léčebná odpověď (LO) (sledovaný jev)	**Zkoumaná léčba** (expozice)	
	ano	ne
neúspěch (LO nedosažena)	80	40
úspěch (LO dosažena)	20	60
odhad rizika (pravděpodobnosti) neúspěchu	80/100 = 0,80	40/100 = 0,40
odhad pravděpodobnosti úspěchu	20/100 = 0,20	60/100 = 0,60
odhad šance – neúspěch	80/20 = 4,00	40/60 = 0,67
odhad šance – úspěch	20/80 = 0,25	60/40 = 1,50

Výpočet č. 1.
Zkoumáme vliv léčby z hlediska rizika (pravděpodobnosti) neúspěchu, resp. šance na neúspěch
- *OR* vypočítáme jako: *OR* = 4,00/0,67 = 6,000
- *RR* vypočítáme jako: *RR* = 0,80/0,40 = 2,000

Výpočet č. 2.
Zkoumáme vliv léčby z hlediska pravděpodobnosti úspěchu, resp. šance na úspěch
- *OR* vypočítáme jako: *OR* = 0,25/1,50 = 0,167
- *RR* vypočítáme jako: *RR* = 0,20/0,60 = 0,333

Výpočet č. 3.
Ověření recipročního přepočtu OR
- Inverzním výpočtem ověříme symetrii hodnoty OR, kterou lehce změníme cíl výpočtu. Platí $OR_{neúspěch} = 1/OR_{úspěch}$, a tedy 6,00 = 1/0,167.
- A naopak $OR_{úspěch} = 1/OR_{neúspěch}$, a tedy 0,167 = 1/6,00.
- Uvedený inverzní výpočet ale neplatí pro hodnotu *RR*, kde inverzní funkcí nepotvrdíme hodnotu opačnou, získanou přímo ze zdrojové tabulky: 1/2,00 = 0,50 ≠ 0,333.

Závěry
- Zkoumaná experimentální léčba nevede ke slibným výsledkům, riziko neúspěchu je 80 %, zatímco ve skupině bez této léčby je to hodnota poloviční. Šance na neúspěch je ve skupině s léčbou 4,0; zatímco ve skupině bez léčby pouze 0,67.
- Zkoumáme běžný, tedy nikoli vzácný jev a v důsledku toho se hodnoty *OR* a *RR* mezi sebou významně liší. Dokumentujme tento fakt na neúspěchu léčby jako cílovém stavu, který dosahuje 40% prevalence v kontrolní skupině. Hodnota *OR* je trojnásobně vyšší a více vzdálená od referenční hodnoty 1,00 než *RR* (výpočet č. 1).
- *OR* i *RR* můžeme vyhodnotit se zaměřením na opačný cílový stav: neúspěch léčby (výpočet č. 1) vs úspěch léčby (výpočet č. 2). Ze srovnání výsledků těchto dvou výstupů vidíme, že jak *OR*, tak *RR* shodně ukazují výrazný rizikový vliv zkoumané léčby (při zaměření na neúspěch jsou hodnoty *OR* a *RR* výrazně větší než 1, při zaměření na úspěch naopak).
- Odhad hodnoty *OR* můžeme pomocí inverzní funkce snadno symetricky „přepínat" mezi cílovými parametry úspěch vs neúspěch léčby. Tento přepočet však nefunguje u odhadu *RR*, kde vždy musíme provádět výpočet přímo ze zdrojové tabulky a musíme tak pozorně dbát na to, jak je analýza nastavena.

Příklad 1. Ukázka recipročního přepočtu poměru šancí a relativního rizika v asociační studii.

Proč je v literatuře stále častěji používán *OR*, a to i u prospektivních kohortových studií?

Důvodů je více. V první řadě můžeme uvést snadný způsob výpočtu a jeho dobrou dostupnost v běžných softwarových nástrojích. Více sofistikovaný důvod je, že odhad *OR* je běžný výstup složitějších modelovacích technik, které jsou velmi efektivní pro hodnocení komplikovaných tabulek četností. Zejména jde o logistickou regresi, které se budeme věnovat v některém z dalších dílů našeho seriálu. Tato modelovací technika umožňuje odhady i vícerozměrně adjustovaného poměru šancí, a tedy odfiltrování současného vlivu více zavádějících faktorů. Právě proto bývá využívána i pro kohortové prospektivní studie.

Odhad *OR* má také řadu dalších numerických výhod, které ho často favorizují před *RR*. Například platí následující vztah:

- $OR_{riziko} = 1/OR_{protekce}$ nebo jiným označením $OR_{neúspěch} = 1/OR_{úspěch}$

Jinými slovy jednoduchou inverzí můžeme přepínat odhad *OR* pro hodnocení neúspěchu nebo úspěchu, a to nejen vlastní bodový odhad *OR*, ale i hraniční hodnoty jeho intervalu spolehlivosti. Výhodou je, že můžeme velmi lehce měnit interpretaci stejného odhadu z pohledu rizikovosti nebo naopak protekce dané sledovaným faktorem. Při publikaci výsledků, kde zpracováváme více odhadů *OR*, pak můžeme snadno nastavit jejich „směr" tak, aby vypovídaly konzistentně. Odhad *RR* takovou vlastnost nemá, a tudíž musíme již při primárním výpočtu velmi pozorně sledovat, zda hodnotíme „úspěch" či „neúspěch", nebo „rizikový vliv faktoru vs. protektivní vliv faktoru". Ukázku

proLékaře.cz | 16.8.2025

Při analýze vlivu většího množství faktorů na výskyt nemoci často nastává situace, kdy některé faktory vedou k hodnotě poměru šancí (*OR*) nad jedna (jde tedy o rizikové faktory, rizikové kategorie) a některé naopak pod jedna (tyto faktory jsou nastaveny jako protektivní). Z hlediska interpretace je nicméně vhodnější, pokud všechny analyzované faktory mají stejný „směr“, tedy buď „rizikový“ nebo „protektivní“. Toho je možno dosáhnout jednoduchou inverzí definice rizikové/protektivní skupiny, kdy výsledný poměr šancí je počítán jako převrácená hodnota původního poměru šancí, totéž platí i pro intervaly spolehlivosti poměru šancí. Tento postup je demonstrován na souboru pacientů popisujícím vliv různých charakteristik na výskyt určitého symptomu jako cílového parametru.

Původní varianta (1)

Faktor		Výskyt symptomu ne	Výskyt symptomu ano	*OR* (95% IS)
muž				
	ne	35	17	1,235
	ano	30	18	(0,542; 2,813)
věk				
	≤ 70	45	15	3,000
	> 70	20	20	(1,280; 7,032)
BMI				
	≤ 25	*30*	*30*	*0,143*
	> 25	*35*	*5*	*(0,049; 0,414)*
hypertenze				
	ne	40	16	1,900
	ano	25	19	(0,827; 4,365)
diabetes mellitus				
	ne	41	14	2,563
	ano	24	21	(1,103; 5,955)

Kategorie BMI > 25 vede na rozdíl od ostatních faktorů k *OR* < 1. Pro dosažení stejného směru *OR* u všech faktorů je třeba zaměnit ve výpočtu *OR* rizikovou kategorii BMI.

$$OR(2) = \frac{1}{OR(1)}$$

Nová varianta (2)

Faktor		Výskyt symptomu ne	Výskyt symptomu ano	*OR* (95% IS)
muž				
	ne	35	17	1,235
	ano	30	18	(0,542; 2,813)
věk				
	≤ 70	45	15	3,000
	> 70	20	20	(1,280; 7,032)
BMI				
	> 25	*35*	*5*	*7,000*
	≤ 25	*30*	*30*	*(2,413; 20,306)*
hypertenze				
	ne	40	16	1,900
	ano	25	19	(0,827; 4,365)
diabetes mellitus				
	ne	41	14	2,563
	ano	24	21	(1,103; 5,955)

Po výměně rizikové kategorie BMI > 25 za kategorii BMI ≤ 25 je směr všech hodnot *OR* v analýze shodný (kategorie faktorů jsou nastaveny jako rizikové).

Příklad 2. Přepočet hodnot poměru šancí (*OR*) z „šance protekce“ na „šanci rizika“.

takových recipročních výpočtů dokumentujeme na příkladu 1.

V příkladu 2 dále popisujeme situaci, kdy hodnotíme asociaci více potenciálních prediktorů vůči cílovému parametru „nemoc: ano/ne“. Hodnotíme tedy více tabulek četností a výsledkem je více odhadů *OR*. Pro čtenáře je vhodné a přehledné, pokud zkoumané prediktivní faktory mají stejný „směr“, tudíž vyjadřují buď rizikový, nebo naopak protektivní účinek. Toto sladění provedeme u odhadů *OR* velmi snadno pomocí výše uvedeného vztahu mezi *OR* pro úspěch a *OR* pro neúspěch. Příklad 2 také dokumentuje, že stejně lze postupovat i u hranic intervalu spolehlivosti.

Je možné jednoduše srovnat hodnoty *RR* a *OR*? A za jakých situací?

Zjednodušeně shrnuto má takové přímé srovnání smysl v podstatě jen ve dvou modelových situacích:

- U retrospektivních studií (typicky studie případů a kontrol) při platnosti předpokladu vzácného onemocnění (viz díl 48 a 49 seriálu), kdy lze hodnotu *OR* aproximovat na *RR*.
- U prospektivních studií (typicky observační kohortové nebo průřezové studie, případně i randomizované klinické studie), kde můžeme současně odhadovat *RR* i *OR*; v takovém případě:
  - hodnoty *OR* a *RR* se absolutně shodují pouze v případě, kdy žádná asociace mezi zkoumaným jevem a expozicí neexistuje, potom platí *OR* = *RR* = 1,
  - hodnoty *OR* jsou vždy extrémnější (tj. vzdálenější od hodnoty 1, a to oběma směry) než odpovídající hodnoty *RR*; tento rozdíl narůstá s rostoucí prevalencí zkoumaného znaku ve zdrojové tabulce četností – u běžně se vyskytujících znaků (prevalence větší než 5–10 %) nelze odhady *RR* a *OR* vzájemně zaměňovat.

V jakém vztahu je odhad rizika a odhad šance?

V literatuře je většinou pojednáváno o srovnání a vzájemné aproximaci *OR* a *RR* (viz odstavec výše). Srovnání šance výskytu jevu *A* (O_A) a rizika jevu *A* (P_A) je ale stejně důležité. Připomeňme, že rizikem zde označujeme pravděpodobnost nastání nějakého jevu či události. Šance nastání jevu a jeho riziko ve smyslu pravděpodobnosti jsou vzájemně spojeny již definičním vztahem:

- $O_A = P_A/(1 - P_A)$, a tedy recipročně $P_A = O_A/(1 + O_A)$.

Z těchto vztahů opět vidíme, že u vzácných jevů, kde je P_A nízké, bude hodnota O_A blízká až rovna hodnotě P_A. U běžných jevů budou hodnoty číselně nesouměřitelné, neboť platí:

- P_A je pravděpodobnost, a tudíž může nabývat hodnot pouze v intervalu 0 až 1;
- naopak O_A nabývá hodnot od 0 až do nekonečna.

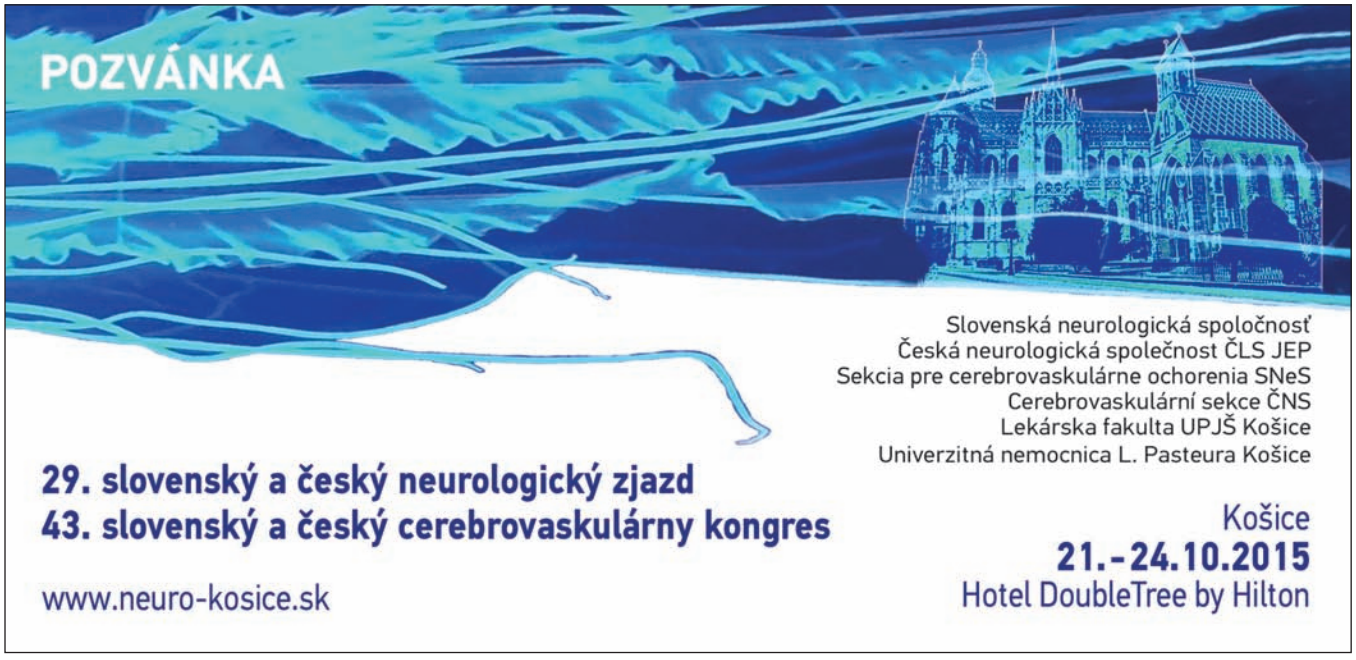
Výše uvedené omezení možných hodnot rizika (P_A) se následně promítá i do omezení rozsahu hodnot poměru rizik, tedy relativního rizika *RR*. Relativní riziko je totiž poměrem rizik, tedy pravděpodobností, a ty jsou li-

proLékaře.cz | 16.8.2025

Tab. 2. Srovnání různých odhadů relativního rizika (*RR*) a poměru šancí (*OR*) z asociačních tabulek četností 2 x 2.
Modelové příklady odvozené z prospektivních kohortových studií, u kterých lze smysluplně hodnotit *RR* i *OR*.

Modelové příklady	Ukazatele vypočítané z tabulky četností (pro ilustraci výpočtu viz příklad 1)			
	riziko v experimentální skupině	riziko v kontrolní skupině	*RR*	*OR*
1) experimentální zásah působí jako rizikový faktor (*RR* > 1)				
příklad 1	0,002	0,001	2,00	2,00
příklad 2	0,200	0,100	2,00	2,25
příklad 3	0,900	0,800	1,13	2,25
příklad 4	0,800	0,400	2,00	6,00
příklad 5	0,900	0,200	4,50	36,00
2) experimentální zásah působí jako protektivní faktor (*RR* < 1) nebo je *RR* = 1				
příklad 6	0,999	0,998	1,00	2,00
příklad 7	0,300	0,700	0,42	0,18
příklad 8	0,200	0,600	0,33	0,17

mitovány rozsahem možných hodnot 0 až 1. Tudíž i hodnoty *RR* jsou omezeny, a proto také hodnoty *OR* vycházejí „extrémněji", tedy jsou číselně vzdálenější od referenční hodnoty 1 než hodnoty *RR*. Uvažujme hodnotu rizika v kontrolní skupině stejnou jako v příkladu 1, tedy 0,4. Potom je maximální možná vypočtená hodnota *RR* = 2,5, protože riziko v čitateli poměru *RR* nemůže překročit 1, a tedy 1/0,4 = 2,5. Hodnota rizika v kontrolní skupině tak determinuje maximální možnou hodnotu *RR*; s klesajícím rizikem v kontrolní skupině roste maximální možná hodnota *RR*. Hodnoty poměru šancí takové omezení nemají. Srovnání různých hodnot *RR* a *OR* uvádíme v tab. 2, ze které je patrné, že:

- s rostoucí prevalencí sledovaného rizikového jevu narůstá rozdíl mezi *OR* a *RR*; *OR* je přitom vždy vzdálenější od hodnoty 1;
- i v případě, kdy je hodnota *RR* téměř rovna 1, může *OR* nabývat hodnoty > 1;
- hodnota *OR* na rozdíl od *RR* souvisí s absolutním rozdílem rizika mezi kontrolní a experimentální skupinou, s rostoucí diferencí rizik roste i *OR*;
- *RR* nesouvisí s absolutním rozdílem rizik; tedy i při stejné absolutní diferenci rizika mezi kontrolní a experimentální skupinou může *RR* nabývat různých hodnot.

Velký číselný rozsah možných hodnot *OR* (tj. od 0 až do nekonečna) představuje problém při vzájemném srovnávání výstupů různých studií. Pro tento účel existují standardizované postupy, díky kterým je interpretace hodnot *OR* snazší. Tomuto problému se budeme věnovat v příštím díle seriálu.