proLékaře.cz | 9.5.2024

Analýza dat v neurologii

XLIX. Vzájemná srovnatelnost a aproximace odhadů relativního rizika a poměru šancí

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

V minulém díle seriálu jsme otevřeli problematiku tzv. předpokladu vzácného onemocnění (*rare disease assumption*), jehož uplatněním jsme schopni aproximovat odhad relativního rizika (*RR*) s pomocí odhadu poměru šancí (*OR*) získaného v retrospektivní observační studii. Předpokladem zde je, že při vzácném výskytu studovaného jevu se odhady *RR* a *OR* numericky přibližují a lze je považovat za numericky srovnatelné. Hlavní přínos je možnost srovnat na bázi hodnot *RR* již publikované výstupy studií provedených s různým experimentálním plánem (např. odhad *RR* z kohortových studií s odhady *OR* ze studií případů a kontrol).

Problém, na který ukázal již příklad v minulém díle, je ale platnost tohoto předpokladu v konkrétní situaci. Předpoklad vzácného onemocnění je totiž jen matematický předpoklad, jehož uplatnitelnost by měla být vždy ověřena. Ačkoli různí autoři poměrně často citují hranici prevalence 10 % a nižší jako dostačující pro platnost tohoto předpokladu, ve skutečnosti nejde v takovém případě o nijak vzácný výskyt jevu. Prevalence 10 % znamená výskyt onemocnění u jednoho člověka z 10 osob, zatímco epidemiologická definice vzácné choroby počítá s výskytem 1 : 2 000. Ve skutečnosti neexistuje žádná exaktně daná hranice pro nezkreslenou aproximaci hodnot *RR* z odhadu *OR*, a hranice 10 % má v literatuře velmi mnoho kritiků, zejm. pokud se používá paušálně a bez ověření.

Zde jistě čtenáře napadne legitimní otázka, jak si tedy má výzkumník nematematik ověřit platnost předpokladu vzácného onemocnění v konkrétní situaci? Odpověď je naštěstí jednoduchá, neboť toto ověření je možné i bez složitých matematických postupů. Stačí vyjít ze vstupních údajů, které při srovnání různých studií musí být známy. Představme si publikaci běžné studie případů a kontrol, kde její čtenář vždy zná:

- primární tabulku četností (kvalitní publikovaná studie musí uvádět velikost vzorku, resp. počet analyzovaných subjektů, osob),
- hodnotu prevalence znaku ve skupině případů i kontrol (tím je myšlen odhad prevalence získaný na konkrétním vzorku studie, vypočítaný z tabulky četností),
- výsledný odhad hodnoty *OR*.

Tyto údaje jsou dostačující pro ověření platnosti předpokladu vzácného onemocnění. Aproximace hodnoty *RR* z takto publikovaného odhadu *OR* neznamená nic jiného než předpoklad, že daný odhad *OR* je v dané situaci numericky blízký až totožný s hodnotou *RR*. A tedy, že platí vztah $OR/RR \approx 1$. Známe-li vstupní tabulku četností, nic nám nebrání v dané situaci pokusně vypočítat i hodnotu *RR* a posoudit tak, zda je aproximace možná, respektive vhodná.

Připomeňme si nejjednodušší tabulku četností 2 × 2 s typickým značením četností v jednotlivých polích *a, b, c, d* (tab. 1).Při velmi vzácném výskytu jevu budou četnosti v polích *a, b* tabulky velmi malé a výpočet hodnoty *RR* se bude přibližně rovnat vztahu: $RR \approx [a/c]/[b/d] \approx a*d/b*c \approx OR$. Stejný jev můžeme vysvětlit i pomocí definičního vztahu pro šanci (*O*), kterou například pro rizikový jev A definujeme takto: $O_A = P_A/(1 - P_A)$, kde P_A je pravděpodobnost nastání jevu A, neboli zde riziko jevu A. Pokud je jev A

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

proLékaře.cz | 9.5.2024

Tzv. předpoklad vzácného onemocnění (*rare disease assumption*) je matematickým předpokladem využívaným v hodnocení retrospektivních asociačních studií případů a kontrol. Jde o předpoklad, že při nízké prevalenci onemocnění lze hodnotu relativního rizika aproximovat pomocí poměru šancí, který se u tohoto typu studií používá pro hodnocení vztahu mezi expozicí a výskytem onemocnění. Řada autorů považuje prevalenci 10 % a méně za dostatečně vzácný výskyt onemocnění, aby tento předpoklad byl naplněn. Jak si však ukážeme v příkladech, nelze hranici 10 % považovat za univerzálně platnou.

Tabulka 1. Studie případů a kontrol při různé prevalenci onemocnění v kontrolní skupině (0,01 % až 10 %): odhady poměru šancí (*OR*) a jejich aproximace na relativní riziko (*RR*). Prevalence událostí ve skupině případů je dvojnásobná oproti kontrolní skupině.

Skupina případů		Kontrolní skupina		Prevalence		*OR*	*RR*
s událostí	bez události	s událostí	bez události	případy	kontroly		
2	9 998	1	9 999	0,02 %	0,01 %	2,00	2,00
10	9 990	5	9 995	0,10 %	0,05 %	2,00	2,00
20	9 980	10	9 990	0,20 %	0,10 %	2,00	2,00
100	9 900	50	9 950	1,00 %	0,50 %	2,01	2,00
200	9 800	100	9 900	2,00 %	1,00 %	2,02	2,00
400	9 600	200	9 800	4,00 %	2,00 %	2,04	2,00
600	9 400	300	9 700	6,00 %	3,00 %	2,06	2,00
800	9 200	400	9 600	8,00 %	4,00 %	2,09	2,00
1 000	9 000	500	9 500	10,00 %	5,00 %	2,11	2,00
1 200	8 800	600	9 400	12,00 %	6,00 %	2,14	2,00
1 400	8 600	700	9 300	14,00 %	7,00 %	2,16	2,00
1 600	8 400	800	9 200	16,00 %	8,00 %	2,19	2,00
1 800	8 200	900	9 100	18,00 %	9,00 %	2,22	2,00
2 000	8 000	1 000	9 000	20,00 %	10,00 %	2,25	2,00

Tabulka 2. Studie případů a kontrol při různé prevalenci onemocnění v kontrolní skupině (0,01 % až 10 %): odhady poměru šancí (*OR*) a jejich aproximace na relativní riziko (*RR*). Prevalence událostí ve skupině případů je šestinásobná oproti kontrolní skupině.

Skupina případů		Kontrolní skupina		Prevalence		*OR*	*RR*
s událostí	bez události	s událostí	bez události	případy	kontroly		
6	9 994	1	9 999	0,06 %	0,01 %	6,00	6,00
30	9 970	5	9 995	0,30 %	0,05 %	6,02	6,00
60	9 940	10	9 990	0,60 %	0,10 %	6,03	6,00
300	9 700	50	9 950	3,00 %	0,50 %	6,15	6,00
600	9 400	100	9 900	6,00 %	1,00 %	6,32	6,00
1 200	8 800	200	9 800	12,00 %	2,00 %	6,68	6,00
1 800	8 200	300	9 700	18,00 %	3,00 %	7,10	6,00
2 400	7 600	400	9 600	24,00 %	4,00 %	7,58	6,00
3 000	7 000	500	9 500	30,00 %	5,00 %	8,14	6,00
3 600	6 400	600	9 400	36,00 %	6,00 %	8,81	6,00
4 200	5 800	700	9 300	42,00 %	7,00 %	9,62	6,00
4 800	5 200	800	9 200	48,00 %	8,00 %	10,62	6,00
5 400	4 600	900	9 100	54,00 %	9,00 %	11,87	6,00
6 000	4 000	1 000	9 000	60,00 %	10,00 %	13,50	6,00

Shrnutí: Uváděná hranice prevalence vzácných onemocnění 10 % neumožňuje plnou srovnatelnost odhadu relativního rizika a poměru šancí, zejm. při vyšší hodnotě relativního rizika (při 10% prevalenci a *RR* = 2,00 je *OR* = 2,25, při *RR* = 6,00 je *OR* = 13,50).

Předpoklad vzácného onemocnění je tak možné uplatnit pouze v případě, kdy je ověřena jeho platnost v konkrétní hodnocené situaci (např. výpočtem).

Příklad 1. Platnost předpokladu vzácného onemocnění při odhadu relativního rizika (*RR*) a poměru šancí (*OR*).

vzácný, pak je P_A velmi malé číslo a platí, že $O_A \approx P_A$. V takovém případě lze tedy na základě hodnoty šance aproximovat pravděpodobnost nastání jevu, a tudíž odhad poměru šancí umožňuje aproximovat hodnotu relativního rizika.

Z tohoto pohledu vypadá verifikace předpokladu vzácného onemocnění relativně snadně, neboť jeho platnost se jeví jako závislá především na relativním výskytu případů mezi exponovanými a neexponovanými jedinci. Realita je ale bohužel složitější a kromě vlastní prevalence onemocnění je vztah mezi *RR* a *OR* ovlivněn i absolutní hodnotou relativního rizika a poměru šancí v dané studii. Právě proto je verifikace předpokladu v konkrétní situaci důležitá, neboť zvláště u vyšších hodnot *OR* nemusí být i při relativně nízké prevalenci aproximace na *RR* obhajitelná.

Tento fakt dokumentuje příklad 1, který výše uvedené ověření vztahu *OR* ≈ *RR* shrnuje na různých výstupech modelových studií. Z příkladu je patrné, že:

- při znalosti primárních dat experimentu můžeme současně snadno vypočítat hodnotu *OR* i *RR* a ověření předpokladu vzácného onemocnění je tak dostupné i matematickým neprofesionálům;
- vztah *OR* a *RR* si lze lehce modelově ověřovat na různých vymyšlených příkladech simulujících experimentální situace;
- zvláště obezřetní musíme být v situaci, kdy je prevalence znaku relativně vysoká (3 % a vyšší) a daná retrospektivní studie vedla k odhadu vysoké nebo naopak nízké hodnoty *OR* (> 2,5; < 0,5);
- při neplatnosti předpokladu vzácného onemocnění poměr šancí neaproximuje dobře odhad relativního rizika; čím je studované onemocnění častější, tím více *OR* nadhodnocuje *RR* (při *RR* > 1) anebo ji více podhodnocuje (při *RR* < 1).

Srovnatelnost a aproximace různých odhadů *OR* a *RR* jsou v současnosti aktuální téma rozebírané v řadě recentních prací (viz literatura uvedená v díle seriálu č. 48). K tomu v nemalé míře přispívá současná renezance observačních studií a také reálný nástup personalizované medicíny. Rostoucí počet prognostických a rizikových markerů, které je nutné respektovat, často znemožňuje efektivní realizaci randomizovaných komparativních studií prostě z toho důvodu, že není možné získat adekvátní počty probandů pro narůstající množství ramen takové studie. Využívání retrospektivních studií tento problém sice částečně řeší, avšak jejich výstupy nemohou kvalitou prospektivní komparativní studie zcela nahradit. Publikované observační výstupy tedy následně vyžadují ve zvýšené míře souhrnné analýzy (meta-analýzy), u kterých se aproximaci odhadů *OR* na *RR* často nelze vyhnout.

I proto můžeme v literatuře nalézt návrhy postupů, jak z odhadu poměru šancí korektně usoudit na hodnotu relativního rizika, jež je pro řadu autorů stále lépe interpretovatelnou mírou rizika. Jeden z takových postupů navržený pro kohortové studie publikovali v roce 1998 Zang a Yu. Postup, ve kterém navrhli výpočet hodnoty

proLékaře.cz | 9.5.2024

Mnoho autorů považuje při posuzování vztahu expozice a onemocnění relativní riziko za jednodušší a vhodnější k interpretaci než poměr šancí. Problémem je nicméně závislost hodnoty relativního rizika na prevalenci onemocnění a tedy požadavek odhadovat je pomocí populačně reprezentativních studií jako jsou například prospektivní kohortové studie. Nicméně i v takových asociačních studiích můžeme pomocí určitých výpočetních technik (např. logistická regrese) přednostně odhadnout poměr šancí. Zang a Yu (1998) navrhli postup pro kohortové studie a výpočet hodnoty *RR* z odhadu *OR* adjustovaného na vliv zavádějících faktorů, a to i při neplatnosti předpokladu vzácného onemocnění. Máme-li v takové studii k dispozici informace o absolutním riziku ve skupině bez expozice a výsledný odhad *OR*, existuje možnost jeho přepočtu na relativní riziko:

Vztah pro přepočet *OR* na *RR*: (Zang a Yu, 1998)

$$RR \approx \frac{OR}{1 - R_C + (R_C \times OR)}$$

, kde *RR* je relativní riziko, *OR* je poměr šancí a R_C je míra rizika ve skupině bez expozice, tedy kontrolní („control“).

Číselný příklad aplikace vztahu:

	Onemocnění			
	ano	**ne**	**% onemocnění**	**šance**
Kohorta s expozicí	375	625	37,5%	0,60
Kohorta bez expozice	125	875	12,5%	0,14

***OR* = 4,2**

Tabulka shrnuje výsledek pozorování s relativním výskytem onemocnění v kontrolní skupině bez expozice 12,5 %, s výslednou hodnotou *OR* = 4,2. Potom přepočet z *OR* na *RR* můžeme provést následovně:

$$RR = \frac{4{,}2}{1 - 0{,}125 + (0{,}125 \times 4{,}2)} = 3{,}0$$

Shrnutí: V případě, že známe riziko (odhad pravděpodobnosti) onemocnění v kontrolní kohortě, můžeme pomocí postupu navrženého autory Zang a Yu provést přepočet poměru šancí na relativní riziko i v případě, že není splněn předpoklad vzácného onemocnění. Nicméně je nutné zdůraznit, že daný postup má své velké limity. Na to upozornili ve své reakci na původní práci McNutt et al v roce 1999. Postup navržený Zangem a Yu je použitelný pouze u asociačních studií, kde zavádějící faktory nemodifikují účinek, a tedy hodnoty *RR* jsou stejné ve všech sub-kohortách daných kategoriemi hodnot zavádějících faktorů.

Příklad 2. Výpočet relativního rizika (*RR*) z poměru šancí (*OR*).

RR z odhadu *OR* adjustovaného na vliv zavádějících faktorů, v zjednodušené podobě přibližujeme v příkladu 2. Autoři řeší fakt, že stále častěji využívané metody vedoucí k odhadu *OR* v prospektivních kohortových studiích (např. logistická regrese) neumožňují korektně usuzovat na hodnotu *RR*. Navržená metoda korekce *OR* je cenná i v tom, že umožňuje aproximaci na *RR* pro běžné jevy s prevalencí > 10 %. Korekce odhadu *RR* je nutná především v oblasti hodnot *OR* > 2,5 anebo *OR* < 0,5. Jakkoli je tento postup lákavý, je třeba jej využívat s obezřetností. Na omezení použití této korekce upozornili ve své reakci na původní návrh autoři McNutt et al v roce 1999. Postup navržený Zangem a Yuem je použitelný pouze u asociačních studií, kde zavádějící faktory nemodifikují účinek, a hodnoty *RR* jsou stejné ve všech subkohortách daných zavádějícími faktory. Problém ovšem je, jak takovou situaci správně identifikovat a doložit, a proto má postup uvedený v příkladu 2 poměrně významná omezení v reálném uplatnění.

Současná literatura a dostupné důkazy se většinou přiklánějí k názoru nepoužívat odhady poměru šancí u studií, které umožňují přímočarý odhad relativního rizika. Tedy zejm. u prospektivních kohortových a průřezových studií. Velmi srozumitelnou formou tento problém rozebírá například Lee, na jehož sdělení roku 1994 uvádíme odkaz v seznamu literatury. Ovšem zdá se, že s chybnou záměnou a interpretací odhadů *OR* a *RR* stále bojuje mnoho původních publikací. Podle řady nedávných přehledů literatury je poměr šancí často chybně interpretován jako relativní riziko anebo je relativní riziko nesprávně kalkulováno u retrospektivních studií, kde jej nelze nezkresleně odhadnout. Například Holcomb et al (2001) uvádějí v průzkumu prací ze dvou časopisů více než 25% chybovost v této oblasti. Granados (1995) upozornil na to, že problematická je i definice *OR* a *RR* v řadě využívaných matematických a klinických odborných slovníků, které chybně oba ukazatele definují nebo je zaměňují. Potom se ovšem nelze divit tomu, že se chyby kumulují. Držet se jednoduchých pravidel a používat poměr šancí především tam, kde nemůžeme odhadnout relativní riziko (tj. zvláště u retrospektivních observačních studií), je asi nejlepší cesta, jak si při publikaci vlastních dat nezkomplikovat cestu k jejich interpretaci.

Literatura

Granados JA. Odds and odds ratio: an odd confusion. Epidemiology 1995; 6(5): 571–572.

Holcomb WL jr, Chaiworapongsa T, Luke DA, Burgdorf KD. An odd measure of risk: use and misuse of the odds ratio. Obstet Gynecol 2001; 98(4): 685–688.

Lee J. Odds ratio or relative risk for cross-sectional data? Int J Epidemiology 1994; 23(1): 201–203.

McNutt LA, Hafner JP, Xue X. Correcting the odds ratio in cohort studies of common outcomes. JAMA 1999; 282(6): 529.

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(19): 1690–1691.