proLékaře.cz | 6.11.2026

Analýza dat v neurologii

XXXVII. Statistické testy srovnávající odhady poměru šancí a relativního rizika

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

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

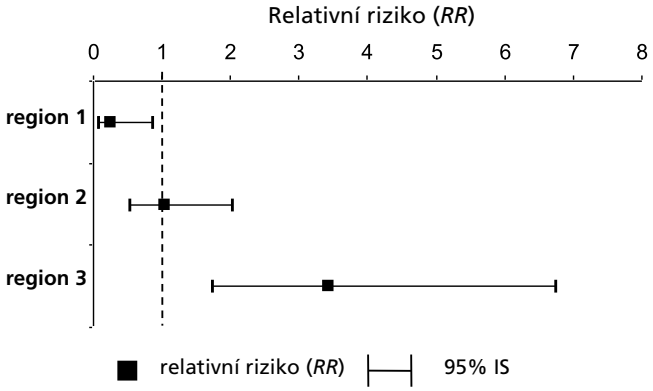
Tento díl statistického seriálu volně navazuje na předchozí části a rozvíjí problematiku odhadu poměru šancí („*odds ratio*", *OR*) a relativního rizika („*relative risk*", *RR*) v tzv. asociačních studiích. Takto označujeme studie (většinou observační, neintervenční), které zkoumají asociaci mezi expozicí nějakým faktorem (rizikovým nebo protektivním) a výskytem nějaké události (vznik nemoci, komplikace nemoci, relaps nemoci, úmrtí apod.); cílem studie je existující asociaci identifikovat a kvantifikovat a následně posoudit její statistickou významnost. Statistické testování je zvlášť důležité, neboť nám pomáhá určit statisticky významné asociace, které nejsou výsledkem náhody a u nichž lze následně hledat odpovídající klinickou nebo epidemiologickou interpretaci. V tomto díle se zaměříme na statistické testy pro odhady *OR* a *RR*, základní vysvětlení jejich

Pacienti s rozdílnými variantami určitého genu (X nebo Y) jsou od narození prospektivně sledováni na výskyt poruch růstu do 15. roku věku. Varianta genu X je zkoumána jako potenciálně riziková pro poruchy růstu. Stejný typ studie je proveden ve třech rozdílných regionech světa se závěrečným vyhodnocením pomocí odhadu relativního rizika (*RR*).

	Region 1			Region 2			Region 3		
Porucha růstu	**gen X**	**gen Y**	**Celkem**	**gen X**	**gen Y**	**Celkem**	**gen X**	**gen Y**	**Celkem**
ano	a = 2	b = 68	a + b = 70	a = 7	b = 28	a + b = 35	a = 14	b = 10	a + b = 24
ne	c = 14	d = 58	c + d = 72	c = 14	d = 58	c + d = 72	c = 14	d = 58	c + d = 72
Celkem	a + c = 16	b + d = 126	142	a + c = 21	b + d = 86	107	a + c = 28	b + d = 68	96

$$RR = \frac{\frac{a}{a+c}}{\frac{b}{b+d}} \qquad RR_1 = \frac{\frac{2}{16}}{\frac{68}{126}} = 0{,}23 \qquad RR_2 = \frac{\frac{7}{21}}{\frac{28}{86}} = 1{,}02 \qquad RR_3 = \frac{\frac{14}{28}}{\frac{10}{68}} = 3{,}40$$

***RR* (95% IS)**	**0,23 (0,06; 0,86)**	**1,02 (0,52; 2,02)**	**3,40 (1,72; 6,72)**



Grafické srovnání ukazuje na velmi podstatné rozdíly ve výstupu jednotlivých studií. Dílčí výsledky se jednak liší ve statistické významnosti jednotlivých odhadů *RR* a dále rovněž mezi sebou. Studie 1 prokázala statisticky významný protektivní vliv genové varianty X, zatímco studie 3 dospěla k opačnému závěru. Výsledky tří studií jsou evidentně nekonzistentní a nelze je společně zobecnit.

Příklad 1. Výpočet a grafické srovnání více odhadů relativního rizika *(RR)*.

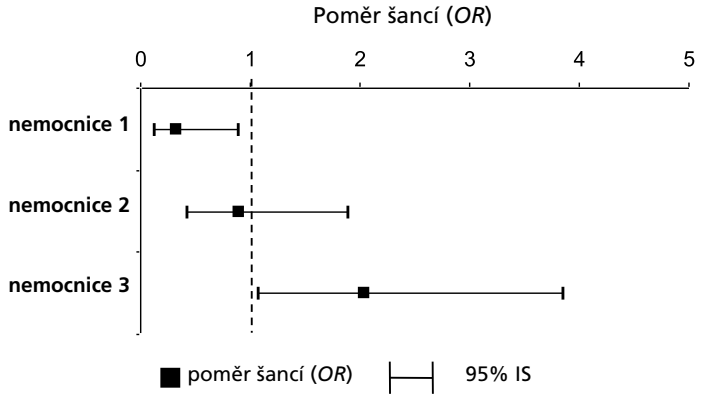
proLékaře.cz | 6.11.2026

U souboru pacientů tří nemocnic byl retrospektivně zjišťován vztah mezi výskytem atopického ekzému a expozicí toxinem zjištěnou dotazováním při vstupním vyšetření. Závěrečné hodnocení bylo založeno na odhadu poměru šancí (odds ratio, *OR*).

	Nemocnice 1			Nemocnice 2			Nemocnice 3		
Ekzém	**expozice toxinem**	**bez expozice**	**Celkem**	**expozice toxinem**	**bez expozice**	**Celkem**	**expozice toxinem**	**bez expozice**	**Celkem**
ano	a = 5	b = 17	a + b = 22	a = 14	b = 17	a + b = 31	a = 32	b = 17	a + b = 49
ne	c = 118	d = 127	c + d = 245	c = 118	d = 127	c + d = 245	c = 118	d = 127	c + d = 245
Celkem	a + c = 123	b + d = 144	267	a + c = 132	b + d = 144	276	a + c = 150	b + d = 144	294

$$OR = \frac{\frac{a}{c}}{\frac{b}{d}} \qquad OR_1 = \frac{\frac{5}{118}}{\frac{17}{127}} = 0{,}32 \qquad OR_2 = \frac{\frac{14}{118}}{\frac{17}{127}} = 0{,}89 \qquad OR_3 = \frac{\frac{32}{118}}{\frac{17}{127}} = 2{,}03$$

***OR* (95% IS)**	**0,32 (0,11; 0,89)**	**0,89 (0,42; 1,88)**	**2,03 (1,07; 3,84)**



Grafické srovnání jednotlivých odhadů *OR* prokázalo významnou heterogenitu v závěrech jednotlivých šetření. Pouze studie v nemocnici 3 prokázala rizikový vliv expozice toxinem, závěry v nemocnici 2 nebyly statisticky významné a zkoumání v nemocnici 1 dokonce indikovalo protektivní vliv expozice toxinem. Závěry studií jsou nekonzistentní a nelze je zobecnit.

Příklad 2. Výpočet a grafické srovnání více odhadů poměru šancí (*OR*).

významu bylo uvedeno v dílech XXXV a XXXVI.

V předchozím výkladu jsme doložili poměrně jednoduchý postup, jak posoudit statistickou významnost jednotlivých odhadů *OR* nebo *RR* pomocí jejich intervalu spolehlivosti. Pokud interval spolehlivosti zahrnuje hodnotu 1, nemůžeme prokázat statisticky významný rozdíl v četnosti výskytu události mezi exponovanými a neexponovanými jedinci, a nemůžeme tudíž prokázat významnou asociaci mezi zkoumanou expozicí a danou událostí. Pokud odhady *OR* či *RR* dokonce dosáhnou přesně hodnoty 1, pak jde o situaci, kdy daný faktor vůbec neovlivňuje výskyt dané události. Pokud však 95% interval spolehlivosti odhadu leží mimo hodnotu 1, potom hovoříme o statistické významnosti odhadu *OR* nebo *RR* na hladině 5 %. Je-li *OR* (*RR*) > 1, pak takto prokazujeme významný rizikový faktor, při *OR* (*RR*) < 1 jde naopak o faktor protektivní. Minulé díly seriálu rovněž zdůraznily, že interval spolehlivosti je třeba považovat za povinný, neboť bez něj nemůže být interpretace odhadů *OR* (*RR*) plnohodnotná.

Jednoduché a jistě snadno pochopitelné využití intervalu spolehlivosti nám takto pomáhá posoudit statistickou významnost jednotlivých odhadů *OR* nebo *RR*. Co ale dělat v případě, kdy potřebujeme vzájemně porovnat více odhadů *OR* či *RR* a otestovat statistickou významnost rozdílů mezi nimi? Tato otázka je zcela na místě, neboť vzájemné srovnávání výstupů asociačních studií je v klinickém výzkumu velmi časté. Uveďme zde několik typických situací, které ke srovnávání odhadů *OR* a *RR* nevyhnutelně vedou:

- Několik vzájemně zcela nezávislých studií případů a kontrol studuje stejné potenciálně rizikové faktory, které mohou způsobovat výskyt závažných komplikací nějaké choroby; chceme porovnat jednotlivé odhady *OR* a zjistit, zda se statisticky významně liší, a zda jsou závěry různých studií konzistentní, či nikoliv.
- U klinické studie pracujeme s kohortou pacientů a zdravých kontrol a hodnotíme asociaci potenciálně rizikových faktorů s výskytem nemoci. Aplikujeme určitý klasifikační (stratifikační) faktor a obě kohorty rozdělíme na dvě nebo i více podskupin (tzv. strata) a srovnáváme, zda je asociace rizikového faktoru s nemocí stejně silná nebo rozdílná v těchto podskupinách.
- Prospektivně získáme ve dvou nemocnicích různé kohorty pacientů, u kterých chceme posoudit vliv vybraných anamnestických faktorů na časný relaps nemoci. Takto separátně odhadnuté hodnoty *RR* je možné vzájemně srovnat, a posoudit tak shodu výsledku v různých kohortách a v případě shody i možnost spojení obou souborů dat.
- V retrospektivní populační studii zkoumáme vliv několika faktorů na vznik ne-

proLékaře.cz | 6.11.2026

U souboru pacientů retrospektivně získaných ve dvou nemocnicích byl zjišťován vztah mezi výskytem ekzému a expozicí toxinem zjištěnou dotazováním při vstupním vyšetření. Cílem příkladu je doložit postup statistického srovnání dvou nezávislých odhadů poměru šancí (*OR*).

	Nemocnice 1			Nemocnice 2		
Ekzém	**expozice toxinem**	**bez expozice**	**Celkem**	**expozice toxinem**	**bez expozice**	**Celkem**
ano	a = 5	b = 17	a + b = 22	a = 14	b = 17	a + b = 31
ne	c = 118	d = 127	c + d = 245	c = 118	d = 127	c + d = 245
Celkem	a + c = 123	b + d = 144	267	a + c = 132	b + d = 144	276

$$OR = \frac{\frac{a}{c}}{\frac{b}{d}} \qquad OR_1 = \frac{\frac{5}{118}}{\frac{17}{127}} = 0{,}32 \qquad OR_2 = \frac{\frac{14}{118}}{\frac{17}{127}} = 0{,}89$$

1.	*OR* (95% IS)	0,32 (0,11; 0,89)	0,89 (0,42; 1,88)
2.	ln(*OR*) (95% IS)	–1,139 (–2,207; –0,117)	–0,117 (–0,868; 0,631)
3.	Výpočet *SE*	$SE_1 = \frac{(-0{,}117 + 2{,}207)}{2*1{,}96} = 0{,}533$	$SE_2 = \frac{(0{,}631 + 0{,}868)}{2*1{,}96} = 0{,}382$
4.	Rozdíl ln(*OR* a jeho *SE*	$\ln(OR_1) - \ln(OR_2) = -1{,}139 + 0{,}117 = -1{,}023$ $SE[\ln(OR_1) - \ln(OR_2)] = \sqrt{SE_1^2 + SE_2^2} = 0{,}656$	
5.	Test významnosti rozdílu ln(*OR*)	$z = \frac{\ln(OR_1) - \ln(OR_2)}{SE[\ln(OR_1) - \ln(OR_2)]} = \frac{-1{,}023}{0{,}656} = -1{,}559$ **=> *p* = 0,115**	
6.	Poměr srovnávaných *OR* a jeho 95% IS	exp(–1,023) [(exp(–1,023 – 1,96 * 0,656); exp(–1,023 + 1,96 * 0,656)] **0,36 (0,10; 1,30)**	

Srovnání dvou odhadů poměru šancí je častou úlohou při statistickém hodnocení. Srovnání dvou odhadů doplněných o střední chybu jejich rozdílu je obecně použitelnou metodou (Altman a Bland, 2003).

Postup výpočtu je následující:
1. Vstupem jsou odhady *OR* a jejich 95% intervaly spolehlivosti.
2. Hodnoty *OR* a jejich IS jsou logaritmicky transformovány pro dosažení normality rozdělení.
3. Z šířky transformovaného intervalu spolehlivosti je odvozena střední chyba obou log(*OR*).
4. Je spočítán rozdíl mezi logaritmicky transformovanými *OR* a jeho střední chyba.
5. Testovou statistiku Z získáme jako poměr rozdílu log(*OR*) a jeho střední chyby; použitým modelem je normální rozdělení.
6. Srovnání obou odhadů *OR* je dále popsáno jejich vzájemným poměrem a intervalem spolehlivosti tohoto poměru.

Výsledná hodnota $p = 0{,}115$ znamená, že nezamítáme nulovou hypotézu shody obou *OR*. Výsledky šetření v obou nemocnicích se na základě dostupných dat statisticky významně neliší.

Příklad 3. Testování statistické významnosti rozdílu mezi dvěma odhady poměrů šancí (*OR*).

moci a chceme vzájemně porovnat jejich „sílu", resp. sílu jejich asociace s cílovou rizikovou událostí (nemoc). Vzájemné statistické srovnání odhadů *OR* umožní jednotlivé faktory lépe diferencovat z hlediska jejich rizikovosti.

Ve všech výše uvedených příkladech lze srovnávané odhady *OR* i *RR* zakreslit do zvláštního typu grafu (tzv. forest plot), který zobrazuje bodové odhady *RR* či *OR* jako body, jejich intervaly spolehlivosti jako horizontální úsečky a také vyznačuje referenční hodnotu 1. Ukázky takových srovnání, tabulkových i grafických, přinášejí příklady 1 a 2. V obou příkladech hodnocení končí konstatováním významné heterogenity výstupů nezávisle provedených studií, z čehož plyne, že jejich závěry z důvodu nekonzistence nelze zobecnit. Použité grafické srovnání odhadů *OR* a *RR* zjednodušuje čtení výsledků, zvláště pokud je srovnávaných dílčích studií větší počet. Dalším formám a aplikacím forest plotu se budeme věnovat v příštím díle seriálu.

Možnost grafického srovnání pomocí intervalů spolehlivosti je lákavá a zjednodušuje práci, ne vždy však jako závěr studie stačí. Zejména posuzování výsledků různých klinických studií velmi často požaduje srovnání různých odhadů *OR* nebo *RR* pomocí statistického testu a posouzení významnosti rozdílu na dané hladině významnosti. Ukázku takového testu pro srovnání dvou odhadů *OR* nebo *RR* přinášejí příklady 3 a 4. Výpočetně se jedná o to, že srovnávané odhady *OR* a *RR* jsou normalizovány logaritmickou transformací, přičemž následný statistický test logicky vyústí ve statistiku se standardizovaným normálním rozdělením (tzv. z-statistiku).

Výpočetní postup srovnání dvou odhadů *OR* nebo *RR* v příkladech 3 a 4 je totožný, neboť jde o poměrové ukazatele odvozené na stejné bázi, tedy ze čtyřpolní tabulky četností. O aplikaci *OR* nebo *RR* totiž nerozhoduje statistický test, ale typ studie: *OR* je odhadnuto v případě retrospektivního sběru dat a *RR* u prospektivní studie. Aplikovaný statistický test také není ničím novým, zcela stejný postup se využívá například při testování rozdílu dvou odhadů střední hodnoty pomocí aritmetického průměru (tzv. t-test pro dva nezávislé výběry). Jde o obecný postup uplatnitelný pro hodnocení statistického významu rozdílů v odhadech všech statistik, které mají normální rozdělení. Rozdíl srovnávaných bodových odhadů příslušné statistiky je vztažen na standardní chybu (*SE*) tohoto rozdílu:

pro poměr šancí (OR):

$$z = \frac{\ln(OR_1) - \ln(OR_2)}{SE[\ln(OR_1) - \ln(OR_2)]}$$

pro relativní riziko (RR):

$$z = \frac{\ln(RR_1) - \ln(RR_2)}{SE[\ln(RR_1) - \ln(RR_2)]}$$

proLékaře.cz | 6.11.2026

Pacienti s potenciálně rizikovou variantou genu X jsou od narození prospektivně sledováni pro výskyt poruch růstu do 15. roku věku, toto sledování je prováděno ve dvou různých regionech. Cílem příkladu je doložit postup statistického srovnání dvou nezávislých odhadů relativního rizika (*RR*).

	Region 1			Region 2		
Porucha růstu	**gen X: ano**	**gen X: ne**	**Celkem**	**gen X: ano**	**gen X: ne**	**Celkem**
ano	a = 2	b = 68	a + b = 70	a = 7	b = 28	a + b = 35
ne	c = 14	d = 58	c + d = 72	c = 14	d = 58	c + d = 72
Celkem	a + c = 16	b + d = 126	142	a + c = 21	b + d = 86	107

$$RR = \frac{\frac{a}{a+c}}{\frac{b}{b+d}} \qquad RR_1 = \frac{\frac{2}{16}}{\frac{68}{126}} = 0{,}23 \qquad RR_2 = \frac{\frac{7}{21}}{\frac{28}{86}} = 1{,}02$$

1. *RR* (95% IS): 0,23 (0,06; 0,86) | 1,02 (0,52; 2,02)

2. ln(*RR*) (95% IS): –1,470 (–2,813; –0,151) | 0,020 (–0,654; 0,703)

3. Výpočet *SE*:

$$SE_1 = \frac{(-0{,}151 + 2{,}813)}{2*1{,}96} = 0{,}679 \qquad SE_2 = \frac{(0{,}703 + 0{,}654)}{2*1{,}96} = 0{,}346$$

4. Rozdíl ln(*RR* a jeho *SE*:

$$\ln(RR_1) - \ln(RR_2) = -1{,}470 - 0{,}020 = -1{,}490$$

$$SE[\ln(RR_1) - \ln(RR_2)] = \sqrt{SE_1^2 + SE_2^2} = 0{,}762$$

5. Test významnosti rozdílu ln(*RR*):

$$z = \frac{\ln(RR_1) - \ln(RR_2)}{SE[\ln(RR_1) - \ln(RR_2)]} = \frac{-1{,}490}{0{,}762} = -1{,}954$$

=> *p* = 0,051

6. Poměr srovnávaných *RR* a jeho 95% IS: exp(–1,490) [exp(–1,490 – 1,96 * 0,762); exp(–1,490 + 1,96 * 0,762)]

0,225 (0,051; 1,004)

Srovnání dvou odhadů relativního rizika je častou úlohou při statistickém hodnocení. Srovnání dvou odhadů doplněných o střední chybu jejich rozdílu je obecně použitelnou metodou (Altman a Bland, 2003).

Postup výpočtu je následující:
1. Vstupem jsou odhady RR a jejich 95% intervaly spolehlivosti.
2. Hodnoty *RR* a jejich IS jsou logaritmicky transformovány pro dosažení normality rozdělení.
3. Z šířky transformovaného intervalu spolehlivosti je odvozena střední chyba obou log(*RR*).
4. Je spočítán rozdíl mezi log(*RR*) a jeho střední chyba.
5. Testovou statistiku Z získáme jako poměr rozdílu log(*RR*) a jeho střední chyby; použitým modelem je normální rozdělení.
6. Srovnání obou odhadů *RR* je dále popsáno jejich vzájemným poměrem a intervalem spolehlivosti tohoto poměru.

Výsledná hodnota *p* = 0,051 znamená, že nezamítáme nulovou hypotézu shody obou *RR*. Nicméně na hladině α = 0,05 jde o hraniční hodnotu a výsledky sledování by tedy měly být interpretovány velmi opatrně.

Příklad 4. Testování statistické významnosti rozdílu mezi dvěma odhady relativního rizika (*RR*).

Z výsledků příkladů 3 a 4 lze dále vyvodit několik velmi důležitých poznatků:

- Vzájemně lze statisticky srovnávat i odhady *OR* a *RR*, které samy o sobě nejsou statisticky významné. I statisticky nevýznamné odhady, tedy takové, jejichž interval spolehlivosti obsahuje hodnotu 1, se mohou vzájemně statisticky významně lišit. A naopak, statisticky významná hodnota *OR* či *RR* se nemusí významně odlišovat od hodnoty, která sama statisticky významná není. To je výstup příkladu 3, kde hodnota OR_1 je statisticky významná, hodnota OR_2 naopak není a oba odhady se navíc statisticky významně neliší ($p = 0{,}115$).
- Výsledek statistického testu je silně závislý na velikosti vzorku a na zastoupení jednotlivých variant v původní tabulce četností. V praxi nejsou výjimkou výpočty, jejichž výsledkem je velmi hraniční zamítnutí nebo naopak potvrzení nulové hypotézy. Takový výsledek zde nabízí příklad 4, kde nezamítáme nulovou hypotézu o rovnosti dvou odhadů *RR* při $p = 0{,}051$. Tento typ výsledku by měl vždy nabádat k velmi opatrné interpretaci s ohledem na možnost zkreslení.

Nespornou výhodou zde doporučeného testu srovnávajícího odhady *OR* nebo *RR* je jeho jednoduchost. Pokud již máme k dispozici provedené odhady *OR* a *RR*, lze výpočet snadno provést i bez pomoci počítače. Snadná aplikovatelnost má ale i své stinné stránky; všimněme si, že do testu lze vstoupit přímo již s bodovým odhadem a jeho intervalem spolehlivosti. Pokud tedy tyto hodnoty máme, například z publikované práce jiného autora, můžeme ihned srovnávat. Pro výpočet testové statistiky nemusíme mít k dispozici původní data z tabulky četností, dokonce ani nemusíme znát velikost souborů, ze kterých jsou srovnávané hodnoty *OR* či *RR* odhadovány. Tyto skutečnosti mohou být zdrojem velmi vážných zkreslení, zvláště při následném srovnávání výsledků již publikovaných prací. Zejména je třeba kontrolovat, zda:

- srovnávané odhady *OR* a *RR* pocházejí ze studií se srovnatelným designem (např. zda obě takto srovnávané studie jsou prospektivní, nebo retrospektivní),
- jsou oba srovnávané odhady *OR* nebo *RR* vzájemně nezávislé, a pocházejí tedy z nezávislých studií,
- zdrojové soubory generující srovnávané odhady *OR* a *RR* nejsou zásadně jinak velké,
- prevalence sledované události není zásadně odlišná v obou souborech dat, ze kterých byly odhadnuty srovnávané hodnoty *OR* nebo *RR*.

Všechny výše uvedené rozdíly mezi srovnávanými studiemi nebo výběrovými populacemi jedinců mohou narušovat věcnou srovnatelnost odhadů *OR* a *RR*. Například pokud bychom postupem v příkladě 3 srovnávali proti sobě

proLékaře.cz | 6.11.2026

dvě hodnoty *OR*, které jsou spočítány z věcně nesrovnatelných studií, je i sebelepší statistický test předem diskvalifikován. Obdobně statistické srovnání hodnot obou ukazatelů odhadnutých ze zcela odlišných výběrů nebo populací může vést k velmi zkresleným závěrům. Přitom ale matematické vztahy umožňují takto nesrovnatelné odhady *OR* a *RR* dosadit a výpočet realizovat! Opět se setkáváme se skutečností, že znalosti a pozornost odborníka, který podstatě zkoumaného problému rozumí, jsou nenahraditelné. Zvláště u asociačních studií je nutné se přesvědčit, zda srovnáváme srovnatelné.

Ačkoliv by se mohlo zdát, že vysvětlením testů v příkladech 3 a 4 problematika srovnávání různých odhadů *OR* a *RR* v našem seriálu končí, zdaleka tomu tak není. Tyto testy totiž řeší pouze srovnání dvou vzájemně nezávislých odhadů, resp. výstupů dvou vzájemně nezávislých asociačních studií. Avšak nemůžeme opomenout vzájemné srovnávání výstupů více než dvou studií, které je v klinickém výzkumu rovněž zcela běžné. Hovoříme o testování homogenity více odhadů *OR* nebo *RR*. Typicky takto posuzujeme míru shody různých mezinárodních studií zaměřených na stejný problém, s cílem jejich výsledky zobecnit (pokud se shodují) anebo rozporovat (pokud jsou statisticky významně rozdílné). Tyto statistické postupy patří do oblasti tzv. metaanalýz klinických studií, které obecně definujeme jako postupy založené na statistické kombinaci výstupů již dříve publikovaných výsledků klinických studií. Věrohodnost těchto „nadstavbových" analýz ovšem zásadně závisí na kvalitě srovnávaných studií a rovněž na správném statistickém srovnávání jejich výsledků. Metodice srovnávání mnohočetných odhadů *OR* a *RR* budeme věnovat příští díly seriálu.

Literatura

Altman DG, Bland JM. Interaction revisited: the difference between two estimates. BMJ 2003; 326(7382): 219.

1. cena časopisu

ČESKÁ A SLOVENSKÁ NEUROLOGIE A NEUROCHIRURGIE

2012

D. Školoudík[1,2], M. Kuliha[1,2], T. Jonszta[3], R. Herzig[1]

[1]Komplexní cerebrovaskulární centrum, Neurologická klinika LF UP a FN Olomouc
[2]Komplexní cerebrovaskulární centrum, Neurologická klinika FN Ostrava
[3]Komplexní cerebrovaskulární centrum, Radiologická klinika FN Ostrava

Endovaskulární léčba ischemické cévní mozkové příhody – minimonografie

Cesk Slov Neurol N 2012; 75/108(6): 669–683