proLékaře.cz | 19.8.2025

Analýza dat v neurologii

LIV. Koncept atributivního rizika v analýze populačních studií – I. Aplikace v různých typech studií

V předchozím díle seriálu jsme jako nové téma otevřeli populační hodnocení rizika v asociačních studiích, kde studujeme vliv expozičního faktoru a jeho následku, např. nemoci. Rozšířili jsme dosud probíraný poměr šancí (*OR*) a relativní riziko (*RR*) o další ukazatele (atributivní riziko, atributivní frakce, atributivní počet), které mohou podstatně rozšířit interpretaci získaných dat. V tomto díle se pokusíme dané téma dále rozpracovat a přiblížit tyto pojmy více pomocí ukázek aplikací.

Nejprve připomeneme nově probírané ukazatele a jejich interpretaci formou jednoduchého číselného příkladu. Zkoumáme v prospektivní kohortové studii vliv rizikového faktoru v anamnéze pacientů na následné komplikace léčby. Číselným výstupem studie je tabulka četností 2 × 2 (tab. 1). K výpočtu všech parametrů lze z této tabulky využít vztahy vysvětlené v předchozím díle seriálu; zde tyto postupy doplňujeme pomocí značení buněk tabulky *a, b, c, d*. Jde o poměrně užitečný nástroj, podle kte-

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

Analyzujeme vliv rizikových faktorů na výskyt nádorového onemocnění z dat prospektivní studie. Vzhledem k různé prevalenci rizikových faktorů v hodnocené populaci je nezbytné posoudit nejen míru jejich rizika, ale i kvantifikovat populační dopad jejich vlivu. K údajům hodnotících rizikovost faktorů je tedy ve výpočtu doplněna populační pravděpodobnost expozice. Výpočet pro uvažované rizikové faktory shrnuje následující tabulka (I_{NE}: incidence v neexponované skupině, I_E: incidence v exponované skupině, *RR*: relativní riziko, *AR*: atributivní riziko, *ARF*: atributivní frakce, I_P: incidence v celé populaci, *PAR*: populační atributivní riziko).

Vznik nádorového onemocnění	Expozice toxické látce v chemickém průmyslu		Kouření		Alkoholizmus		Obezita	
	ne	ano	ne	ano	ne	ano	ne	ano
ne	1 000	20	1 000	100	1 000	200	500	1 000
ano	10	10	10	8	10	8	10	36
incidence	I_{NE} = 10/(1 000 + 10) = 0,010	I_E = 10/(20 + 10) = 0,333	I_{NE} = 10/(1 000 + 10) = 0,010	I_E = 8/(100 + 8) = 0,074	I_{NE} = 10/(1 000 + 10) = 0,010	I_E = 8/(200 + 8) = 0,038	I_{NE} = 10/(500 + 10) = 0,020	I_E = 36/(1 000 + 36) = 0,035
RR	0,333/0,010 = 33,7		0,074/0,010 = 7,5		0,038/0,010 = 3,9		0,035/0,020 = 1,8	
$AR = I_E - I_{NE}$	0,333 – 0,010 = 0,323		0,074 – 0,010 = 0,064		0,038 – 0,010 = 0,029		0,035 – 0,020 = 0,015	
$ARF(\%) = (AR/I_E) \times 100$	(0,323/0,333) × 100 = 97,0 %		(0,064/0,074) × 100 = 86,6 %		(0,029/0,038) × 100 = 74,3 %		(0,015/0,035) × 100 = 43,6 %	
Expozice v populaci (externí data)	99,997 %	0,003 %	70,0 %	30,0 %	85,0 %	15,0 %	75,0 %	25,0 %
Incidence v populaci	I_{NE} = 0,00990	I_P = 0,00991	I_{NE} = 0,010	I_P = 0,029	I_{NE} = 0,010	I_P = 0,014	I_{NE} = 0,020	I_P = 0,023
$PAR = I_P - I_{NE}$	0,00991 – 0,00990 = 0,00001		0,029 – 0,010 = 0,019		0,0142 – 0,0099 = 0,0043		0,0234 – 0,0196 = 0,0038	
$PAF(\%) = (PAR/I_P) \times 100$	(0,00001/0,00991) × 100 = 0,098 %		(0,019/0,029) × 100 = 66,0 %		(0,004/0,014) × 100 = 30,2 %		(0,004/0,023) × 100 = 16,2 %	

Jak vyplývá z hodnot *RR*, *AR* a *ARF*, z hlediska rizikovosti je nejvýznamnější expozice v chemickém průmyslu následovaná kouřením, alkoholizmem a obezitou. Pokud ale zohledníme pravděpodobnost expozice v populaci, pak je kvantitativně nejvýznamnějším faktorem kouření následované alkoholizmem, obezitou a až na posledním místě expozice toxické látce v chemickém průmyslu.

Příklad 1. Interpretační význam atributivního rizika (*AR*) a populačního atributivního rizika (*PAR*) v prospektivní studii.

proLékaře.cz | 19.8.2025

rého lze výpočty rychle technicky provádět a snadno realizovat i bez znalosti matematického aparátu.

Nejprve odhadneme incidenci jevu (komplikace léčby) ve skupině exponované faktorem (I_E) a neexponované faktorem (I_{NE}):

- $I_E = a/(a + b) = 700/1\,000 = 0{,}7$;
- $I_{NE} = c/(c + d) = 300/1\,000 = 0{,}3$.

Hodnoty incidence rizikového jevu také někdy nazýváme rizikem (*R*), zde tedy jde o riziko komplikací v exponované (R_E) a neexponované (R_{NE}) skupině.

Hodnotu incidence rizikového jevu v celém souboru, tedy odhad incidence v celé populaci (I_P), spočítáme takto:

- $I_P = (a + c)/(a + b + c + d) = 0{,}5$.

V dalším kroku odhadneme relativní riziko (*RR*), atributivní riziko (*AR*), atributivní frakci vyjádřenou v procentech (*ARF%*), populační atributivní riziko (*PAR*) a populační atributivní frakci vyjádřenou v procentech (*PAF%*):

Tab. 1. Výsledky kohortové studie zkoumající vliv určitého rizikového faktoru v anamnéze pacientů na komplikace určité léčby*.

Rizikový faktor v anamnéze	Komplikace léčby	
	ANO	NE
ANO	$a = 700$	$b = 300$
NE	$c = 300$	$d = 700$

* Označením *a, b, c, d* se typicky označují buňky čtyřpolní tabulky četností a platí $N = a + b + c + d$.

- $RR = R_E/R_{NE} = [a/(a + b)]/[c/(c + d)] = 0{,}7/0{,}3 = 2{,}33$;
- $AR = R_E - R_{NE} = [a/(a + b)] - [c/(c + d)] = 0{,}7 - 0{,}3 = 0{,}4$;
- $PAR = I_P - I_{NE} = [(a + c)/N] - [c/(c + d)] = 0{,}5 - 0{,}3 = 0{,}2$;
- $ARF\% = [(R_E - R_{NE})/R_E] \times 100 = [(0{,}7 - 0{,}3)/0{,}7] \times 100 = 57{,}1$ %;
- $PAF\% = [(I_P - I_{NE})/I_P] \times 100 = [(0{,}5 - 0{,}3)/0{,}5] \times 100 = 40$ %.

Ve shodě s předchozím dílem tohoto seriálu zde ještě nabízíme ukázku alternativních vztahů využívajících znalosti hodnoty *RR*. Jde pouze o alternativní výpočet, který musí samozřejmě vést ke stejným výsledkům. Tyto vztahy uvádíme zejména proto, že se v literatuře objevují často bez dalších vysvětlení, a mohou tak působit zmatečně:

- $ARF = (RR - 1)/RR = (2{,}33 - 1)/2{,}33 = 0{,}571$;
- $PAF = P_E(RR - 1)/[1 + P_E(RR - 1)] = (0{,}5 \times 1{,}33)/[1 + (0{,}5 \times 1{,}33)] = 0{,}4$;

kde P_E je prevalence rizikového faktoru v celém vzorku (populaci), a tedy $P_E = 1\,000/2\,000 = 0{,}5$.

U retrospektivní studie je incidence onemocnění v získané tabulce četností dána přímo výběrem subjektů a nelze ji tedy objektivně kvantifikovat. V takové situaci je možno využít pouze odhad atributivních frakcí *ARF(%)* a *PAF(%)* a namísto vztahu mezi atributivním rizikem a relativním rizikem (*RR*) je využit vztah k *OR* (poměr šancí). V příkladu hodnotíme vztah mezi kouřením a karcinomem plic.

Karcinom plic	**Kouření**	
	ne	**ano**
ne	1 700	300
ano	750	750

Prevalence kouření v populaci byla zjištěna z externího zdroje a dosahuje hodnoty $P_E = 0{,}4$ (40 %).

$$OR = \frac{\frac{750}{300}}{\frac{750}{1700}} = 5{,}67$$

→ Šance na vznik karcinomu plic je u kuřáků 5,67× větší než ve skupině nekuřáků.

$$ARF(\%) = \frac{OR - 1}{OR} \times 100 = \frac{5{,}67 - 1}{5{,}67} = 82{,}4\%$$

→ Celkem 82,4 % karcinomů plic lze pokládat za důsledek kouření.

$$PAF(\%) = \frac{P_E \times (OR - 1)}{P_E \times (OR - 1) + 1} \times 100 = \frac{0{,}4 \times (5{,}67 - 1)}{0{,}4 \times (5{,}67 - 1) + 1} \times 100 = 65{,}1\%$$

→ V případě eliminace rizikového faktoru kouření lze předpokládat pokles výskytu karcinomu plic o 65,1 %.

Závěr: odhady atributivních frakcí jsou využitelné i u retrospektivních studií za předpokladu zohlednění poměru šancí (*OR*) ve výpočtu namísto odhadu relativního rizika (*RR*).

Příklad 2. Atributivní frakce (*ARF%*) a populační atributivní frakce (*PAF%*) v retrospektivní studii.

proLékaře.cz | 19.8.2025

Stejně jako jiné popisné statistiky i bodové odhady *AR* a *PAR* lze doplnit intervalem spolehlivosti. Výpočet intervalu spolehlivosti pro tyto parametry je postaven na permutačních metodách jako je jacknife a bootstrap (Benichou 1991, Benichou a Gail 1990, Kahn et al 2000, Miettinen 1985, Walter 1978). K výpočtu lze využít např. knihovnu attribrisk v software R, modul epitab, somersd v software Stata nebo software StatsDirect.

V příkladu je analyzován vliv rizikového faktoru kouření na výskyt určitého typu nádorového onemocnění jak v prospektivní studii, tak na úrovni populace se známou prevalencí rizikového faktoru.

Vznik nádorového onemocnění	**Kouření**		
	ne	**ano**	
ne	1 000	100	
ano	10	8	
incidence	I_{NE} = 10/(1 000 + 10) = 0,010	I_E = 8/(100 + 8) = 0,074	
RR	0,074/0,010 = 7,5		
$AR = I_E - I_{NE}$	0,074 – 0,010 = 0,064		→ 95% IS: 0,027–0,130
$ARF(\%) = (AR/I_E) \times 100$	(0,064/0,074) × 100 = 86,6 %		→ 95% IS: 36,5–175,7 %
Expozice v populaci (externí data)	70,0 %	30,0 %	
Incidence v populaci	I_{NE} = 0,010	I_P = 0,029	
$PAR = I_P - I_{NE}$	0,029 – 0,010 = 0,019		→ 95% IS: 0,012–0,026
$PAF(\%) = (PAR/I_P) \times 100$	(0,019 / 0,029) × 100 = 66,0 %		→ 95% IS: 40,9–91,2 %

Příklad 3. Interval spolehlivosti pro atributivní riziko (*AR*) a populační atributivní riziko (*PAR*).

Pro praxi není ovšem rozhodující forma výpočtu, ale především správná interpretace výsledků. Zásadní závěry výše uvedených výsledků shrnujeme v následujících bodech, přičemž interpretaci vždy řadíme ke konkrétním parametrům:

- *R*: pacienti s pozitivní anamnézou mají vyšší riziko výskytu komplikací než osoby bez rizikového faktoru v anamnéze (0,7 vs. 0,3);
- *RR*: pacienti s pozitivní anamnézou mají 2,33× vyšší riziko výskytu komplikací než osoby bez rizikového faktoru v anamnéze;
- *AR*: na 1 000 pacientů s pozitivní anamnézou připadá 400 komplikací vzniklých v důsledku rizikového faktoru v anamnéze;
- *ARF%*: ze všech komplikací v souboru pacientů s pozitivní anamnézou jich 57,1 % bylo způsobeno působením rizikového faktoru, tedy 57,1 % ze 700 komplikací je 400;
- *PAR*: na 1 000 lidí v celém souboru (populaci) připadá 200 komplikací vzniklých v důsledku výskytu rizikového faktoru v anamnéze;
- *PAF%*: ze všech komplikací v celém souboru (v celé populaci) jich 40 % je způsobeno vlivem rizikového faktoru v anamnéze.

Odhad *AR* nachází uplatnění v různých typech studií. Výše uvedený příklad pracuje s daty prospektivní observační studie, ve které se dostáváme k věrohodným odhadům incidence jevu v exponované i v neexponované skupině. U těchto studií jsme schopni objektivně kvantifikovat incidenci sledovaných jevů, a tudíž také exaktně kvantifikovat populační dopad zkoumaných expozičních faktorů. Ve výše uvedeném případě byli jedinci nabráni podle faktoru v anamnéze (řádky tabulky) a odhad výskytu komplikací byl proveden následným sledováním v čase. Odhad incidence tak není zkreslen vlastním náběrovým plánem. U těchto studií lze objektivně odhadnout *RR* a všechny výše uvedené parametry, vč. *AR* a *PAR*.

Čtenáři mohou vnést legitimní námitku, že při analýze relativně jednoduché tabulky četností není nutné kalkulovat tak vysoký počet statistických ukazatelů. Samozřejmě je i zde určující cíl výzkumu a požadovaná interpretace výsledků, uvedené parametry nejsou povinné a nemusí být vždy využity všechny. Nicméně souběžný odhad více ukazatelů je pro výstupy prospektivních studií velkým přínosem, jak dále dokumentujeme v příkladu 1. Zpracováváme v něm sledování expozice pracovníků chemického provozu toxickou sloučeninou, která kauzálně způsobuje jistý druh zhoubného nádoru. Naším cílem je nejen posoudit sílu vztahu mezi tímto faktorem a danou nemocí, ale také kvantifikovat skutečný populační efekt dané expozice. Expozici danou látkou potom srovnáváme s dalšími, významně častějšími rizikovými faktory, jako je kouření, alkoholizmus a obezita. Silný kauzální vztah mezi expozicí toxickou látkou a nemocí vede k velmi vysokým hodnotám *RR*. Srovnáme-li odhady *RR* toxické látky s jinými, méně specifickými, rizikovými faktory daného typu malignity, např. s vlivem kouření, alkoholu nebo obezity, jsou až řádově vyšší. Avšak nespecifické rizikové faktory se týkají mnohem větší populace (mají větší prevalenci) a z hlediska populačního významu tak generují mnohem větší počet nemocných než toxická látka, která exponuje pouze malou populaci profesionálů chemického průmyslu. Doplníme-li odhad relativního rizika atributivním rizikem a zejména kvantifikací pomocí *PAR*, dostáváme komplexnější obraz významu sledované expozice.

proLékaře.cz | 19.8.2025

Přesuňme nyní pozornost od prospektivních studií ke studiím retrospektivním. Ukazatele typu *AR* lze s jistým omezením využít i tady, např. u studií případů a kontrol, ve kterých nemůžeme přímo odhadovat incidenci jevů. Náběr jedinců zde totiž provádíme právě podle toho, zda u nich daný jev nastal nebo nenastal, a odhad *RR* tak nemá smysl. V takovém případě odhadneme hodnotu relativního rizika přepočtem či aproximací z poměru šancí (*OR*) a z výše uvedených parametrů lze vypočítat pouze *ARF%* nebo *PAF%* (Cole, MacMahon, 1971). V příkladu 2 dokumentujeme takový odhad *ARF%* a *PAF%* na učebnicovém příkladu retrospektivní studie zkoumající vliv kouření na výskyt karcinomu plic.

Na závěr tohoto dílu doplníme dané téma příkladem 3, který ukazuje výpočet intervalu spolehlivosti pro odhady *AR* a *PAR*. Tak jako např. *RR* lze i *AR* doplnit hranicemi spolehlivosti, které usnadňují jeho interpretaci. Výpočet je v tomto případě bohužel složitější a je založen na permutačních metodách, kterým se budeme v našem seriálu věnovat později. Příklad 3 dále shrnuje pro případné zájemce podstatné zdroje literatury. K výpočtu je nutné využít specializovaný software, např. knihovnu attribrisk v software R, modul epitab, somersd v software Stata nebo software StatsDirect. Zájemcům o tyto výpočty doporučujeme velmi kvalitní a hojně citovanou práci Mezzettiho et al z roku 1996.

Literatura

Benichou J. Methods of adjustment for estimating the attributable risk in case-control studies. A Rev Stat Med 1991; 10(11): 1753–1773.

Benichou J, Gail M. Variance calculations and confidence intervals for estimates of the attributable risk based on logistic models. Biometrics 1990; 46(4): 991–1003.

Cole P, MacMahon B. Attributable risk percent in case-control studies. Br J Prev Soc Med 1971; 25(4): 242–244.

Kahn M, O'Fallon W, Sicks J. Technical report series No. 54, generalized population attributable risk estimation. Rochester, Minnesota: Department of Health Science Research, Mayo Clinic 2000.

Miettinen OS, Nurminen M. Comparative analysis of two rates. Stat Med 1985; 4(2): 213–226.

Walter S.D. Calculation of attributable risks from epidemiologic data. Int J Epidemiol 1978; 7(2): 175–182.

Mezzetti M, Ferraroni M, Decarli A, La Vecchia C, Benichou J. Software for attributable risk and confidence interval estimation in case-control studies. Comput Biomed Res 1996; 29(1): 63–75.

RESOURCE®
ThickenUp® Clear

Sypké zahušťovadlo z xantanové gumy k zahuštění tekutin a pokrmů

- jednoduchá a rychlá příprava
- stabilní konzistence i po delší době
- nemění chuť zahuštěných tekutin nebo pokrmů

Dietní potravina pro zvláštní lékařské účely.

Pro snadnější a bezpečnější polykání

www.nestlehealthscience.cz

NestlēHealthScience