proLékaře.cz | 22.8.2025

Analýza dat v neurologii

LXII. Pozor na interpretaci ekologických (korelačních) studií – I.

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

V tomto díle seriálu rozšíříme výklad výpočtů prováděných na základě tabulek četností, tedy výpočtů směřujícím k odhadům poměru šancí (*„odds ratio"*) a relativního rizika (*„relative risk"*). Oběma pojmům jsme věnovali mnoho předchozích dílů a věříme, že se nám podařilo čtenáře přesvědčit, že jsou v odborné literatuře zcela zdomácnělé a lze je již považovat za přirozenou součást biomedicínské terminologie.

Připomeňme, že odhady poměru šancí a relativního rizika se spolu setkávají zejména v epidemiologických studiích, které zkoumají asociaci mezi rizikovými či ochrannými faktory a výskytem nemocí. Zjišťujeme, zda přítomnost určitého faktoru zvyšuje riziko vzniku nemoci, nebo naopak jejímu vzniku zabraňuje. Obecně tento typ problému vede k vytvoření tabulky četností 2 × 2, kde zkoumáme vztah (asociaci) mezi expozicí daným faktorem a výskytem sledované změny zdravotního stavu. V klinické praxi jsou typické asociační tabulky, do kterých zadáváme přímo individuální data jednotlivých zkoumaných osob, např. kouření (ano/ne) vs. dýchací problémy (ano/ne), vždy u konkrétního jedince. V tomto díle se budeme věnovat zvláštnímu typu takových studií, které nazýváme **studie ekologické (korelační)**. I zde zkoumáme vztah mezi předpokládanou expozicí a zdravotním stavem, avšak nikoli individuálně, ale mezi definovanými skupinami (kohortami) osob. Základní odlišností od „běžných" klinických pozorování je fakt, že zde předmětem analýzy není jedinec, ale celá skupina. Termín „ekologická studie" je zde tedy jakýmsi ekvivalentem výrazu „skupinová studie". Do tabulek četností tak není zaznamenávána pozice jednotlivých osob (kontrol, exponovaných jedinců), ale celých skupin, které jsou definovány např. geograficky (souběžně sledované populace v různých regionech) nebo časově (sledování vývoje v čase v jedné skupině). Místo individuálních záznamů zde pracujeme s agregovanými a skupinovými proměnnými, které vyjadřují situaci v celé kohortě.

Tento postup je často objektivně vynucen u problémů, kde individuální data nejsou k dispozici nebo je z podstaty problému nelze zjistit. Uveďme několik příkladů takových situací:

- Studie zaměřené na environmentální problémy, např. zkoumání vlivu expozice škodlivinami z ovzduší na porodní hmotnost novorozenců, kterou zkoumáme na populačních skupinách několika obcí nebo regionů. Typicky jde o observační retrospektivní studie. V takových situacích samozřejmě nemáme data o každém novorozenci a nejsme schopni přesně kvantifikovat individuální úroveň expozice.
- Jiným příkladem mohou být tzv. mortalitní mapy kvantifikující regionální mortalitu z příčiny určité choroby (např. zhoubné nádory) ve vztahu k různým faktorům prostředí, např. k expozici radonem. Takto byl např. úspěšně zmapován vliv expozice slunečnímu UVB záření na narůstající incidenci zhoubného melanomu. Geografická srovnání vztahující incidenci či mortalitu na určitou chorobu k prevalenci rizikových faktorů jsou jedním z nejčastějších příkladů ekologických studií.
- Studie zkoumající vliv výživy či faktorů životního stylu na vznik a vývoj různých nemocí jsou také velmi častým příkladem ekologických zkoumání, zejména v mezinárodním srovnání. Jde o práce srovnávající stravovací a životní návyky populací různých států či oblastí a vztahující tyto faktory k určitému typu nemocnosti a k výskytu zdravotních problémů.
- Data sledující charakteristiky migrantů, změny jejich životního stylu po migraci a související zdravotní faktory představují velmi cenný vstup do hodnotných ekologických studií. Taková sledování umožňují částečně odlišit vliv genetických faktorů migrující populace od environmentálních a dalších externích vlivů, které byly migrací změněny. I zde je však nutná velká opatrnost při zobecnění závěrů, neboť migranti nemusí být reprezentativním vzorkem původní populace a mohou sebou nést řadu specifických charakteristik.

Ekologické studie nemusí vždy srovnávat pouze různé geograficky vymezené populace, časté je i srovnávání časových period. Může jít např. o hodnocení vlivu stravovacích zvyklostí na nemocnost lidí narozených v minulém století vs. recentně narozená populace; vlivu polychlorovaných bifenylů a dioxinů na zdravotní stav populace před a po řízené regulaci jejich použití (tzv. Stockholmská úmluva o persistentních organických polutantech z roku 2001), apod. Srovnáváme-li takto vývoj a zdravotní stav populací narozených v různých obdobích, hovoříme o tzv. efektu kohorty dle narození (***„birth-cohort effect"***). I zde musí být interpretace velmi opatrná s ohledem na množství potenciálně zavádějících faktorů. Například srovnání dvou kohort narozených v různých obdobích může ve vyspělých státech odhalit narůstající prevalenci řady chorob. Zdánlivě rostoucí riziko však nemusí vůbec souviset s faktory daného období, ale s faktem, že se významně prodlužuje střední doba dožití v těchto zemích. Tento jev je logicky pozorovatelný zejména u chorob s incidenčním výskytem ve vyšším věku, např. diabetes mellitus 2. typu, zhoubné nádory nebo věkem podmíněná makulární degenerace.

Prodlužující se doba života, inovace zdravotní péče, její rostoucí dostupnost a účinnost jsou velmi významnými kofaktory, které limitují populační srovnání z různých časových period. Riziko dezinterpretace při srovnávání různých časových období můžeme významně snížit plánováním dlouhodobých sledování u jasně definované kohorty osob.

proLékaře.cz | 22.8.2025

Tyto dlouhodobé kohortové studie pak umožňují srovnávat různé věkové kategorie na stále stejné skupině probandů. Dlouhodobé kohortové studie mohou být i vícegenerační, což umožní srovnání zdravotního stavu rodičů a dětí. V takovém případě jde ovšem o projekty časově i finančně velmi náročné, a tudíž stále relativně vzácné v mezinárodní literatuře. Nicméně, je-li již taková kohorta založena, je snahou experimentátorů sledovat co nejširší spektrum faktorů, od environmentálních vlivů, sociálních charakteristik, až po faktory zdravotní a genetické. Pro takto široce pojaté projekty sledující v čase působení komplexní škály vnějších faktorů (vč. životního stylu) na lidské zdraví se nově používá pojem expozomové studie (***„exposome studies“***). Široký záběr sledování a zejména longitudinální záznamy u stále stejné kohorty dávají těmto studiím šanci odhalit vliv dosud neznámých rizikových faktorů a jejich případné synergie. To je cenné zejména u chronických a komplexních poruch zdraví, kde sledování samotných genetických faktorů nepřináší dostatečné znalosti k vysvětlení k jejich vzniku (Wild 2005; Rappaport a Smith 2010; Wild 2012).

Z výše uvedených příkladů pro čtenáře jistě vyplývají všechny významné limity ekologických studií. Jde o sledování, u nichž je expozice rizikovým faktorem charakterizována agregovanou expozicí celé skupiny osob (např. při srovnávání různých regionů či zemí). Problém potom nespočívá pouze v samotné heterogenitě takto rozsáhlých populací, ale také v objektivním faktu, že jedinci postižení studovanou nemocí vůbec nemusí být přímo exponováni uvažovanému riziku. Ve skupinových sledováních nemáme pod kontrolou individuální expozici a nekontrolujeme ani pozaďové riziko sledovaných rizikových jevů. Říkáme, že tyto studie jsou vystaveny specifické formě chyb nazývané ekologické zkreslení, někdy též ekologický klam (***„ecological fallacy“***). Tento pojem jednoduše vyjadřuje fakt, že korelace pozorovaná na skupinové (populační) úrovni nemusí vůbec existovat na úrovni individuální.

Ekologické zkreslení limituje uplatnění ekologických sledování zejména při průkazu kauzality vztahů (Piantadosi et al 1988).

Při skupinových srovnáních dále čelíme vlivu neznámých rizikových faktorů nebo tzv. matoucích (***„confounding“***) faktorů, které mohou vést k podstatným zkreslením závěrů. Rovněž nevyváženost různých podskupin může vést ke zkreslením typu Simpsonova paradoxu (Altman a Deeks 2002; Cates 2002; viz též díly 41–42 tohoto seriálu).

Avšak i přes všechny velmi významné limitace a problémy mají ekologická sledování své pevné místo ve studiu vlivu rizikových faktorů na vznik a vývoj nemocí. O tom ostatně svědčí rostoucí počet citací v současné epidemiologické odborné literatuře. Pokud jsou si autoři vědomi rizik zkreslení a seriózně je ve své práci uvedou, nelze proti ekologickým sledováním nic namítat; jako příklad takto korektně zpracované publikace doporučujeme práci Granta z roku 2014, která pojem ekologické sledování obsahuje již v samotném názvu. Takovéto práce mohou významně přispět k pochopení mechanizmů účinku různých rizikových faktorů v reálném světě, zejména pak mohou definovat nové hypotézy pro další výzkum. Řada ekologických studií je základem pro další, více specializované, experimenty (ideálně randomizované prospektivní studie), které exaktněji sledují vliv vybraných faktorů na zkoumané jevy. Takto byla i v minulosti definována řada kauzálních vztahů, o kterých dnes již nepochybujeme. Zmiňme závěrem jako příklady práce zaměřené na expozici UVB záření ve vztahu k incidenci nádorových onemocnění (Gant 2009), studie zaměřené na zkoumání vlivu infekcí na riziko CMP (Grau et al 2010) či velmi četné výzkumy vztahující faktory životního stylu k riziku obezity a dalších civilizačních chorob (Wild 2005; Hu 2013).

Není sporu o tom, že ekologické studie jsou dnešní mezinárodní vědeckou komunitou akceptovány jako důležitý nástroj pro nastavování nových témat a formování hypotéz, avšak nemohou sloužit jako exaktní nástroj prokazující příčinné souvislosti. Vzhledem k významu tohoto tématu se budeme ekologickým studiím v souvislosti s kauzalitou vztahů věnovat ještě v příštím díle našeho seriálu.

Literatura

Altman DG, Deeks JJ. Meta-analysis, Simpson's paradox, and the number needed to treat. BMC Med Res Methodol 2002;2:3.

Cates CJ. Simpson's paradox and calculation of number needed to treat from meta-analysis. BMC Med Res Methodol 2002;2:1.

Davey SG, Phillips A. Confounding in epidemiological studies: why "independent" effects may not be all they seem. BMJ 1992;305(6856):757–9.

Grant WB. How strong is the evidence that solar ultraviolet B and vitamin D reduce the risk of cancer? An examination using Hill's criteria for causality. Dermatoendocrinol 2009;1(1):17–24.

Grant WB. A multicountry ecological study of cancer incidence rates in 2008 with respect to various risk-modifying factors. Nutrients 2014;6(1):163–89. doi: 10.3390/nu6010163.

Grau AJ, Urbanek C, Palm F. Common infections and the risk of stroke. Nat Rev Neurol 2010;6:681–94.

Hu FB. Resolved: there is sufficient scientific evidence that decreasing sugar-sweetened beverage consumption will reduce the prevalence of obesity and obesity-related diseases. Obesity Reviews 2013;14(8):606–19. doi: 10.1111/obr.12040.

Piantadosi S, Byar DP, Green SB. The ecological fallacy. Am J Public Health 1988;127:893–904.

Rappaport SM, Smith MT. Epidemiology. Environment and disease risks. Science 2010;330(6003):460–1. doi: 10.1126/science.1192603.

Wild CP. The exposome: from concept to utility. Int J Epidemiol 2012;41(1):24–32. doi: 10.1093/ije/dyr236.