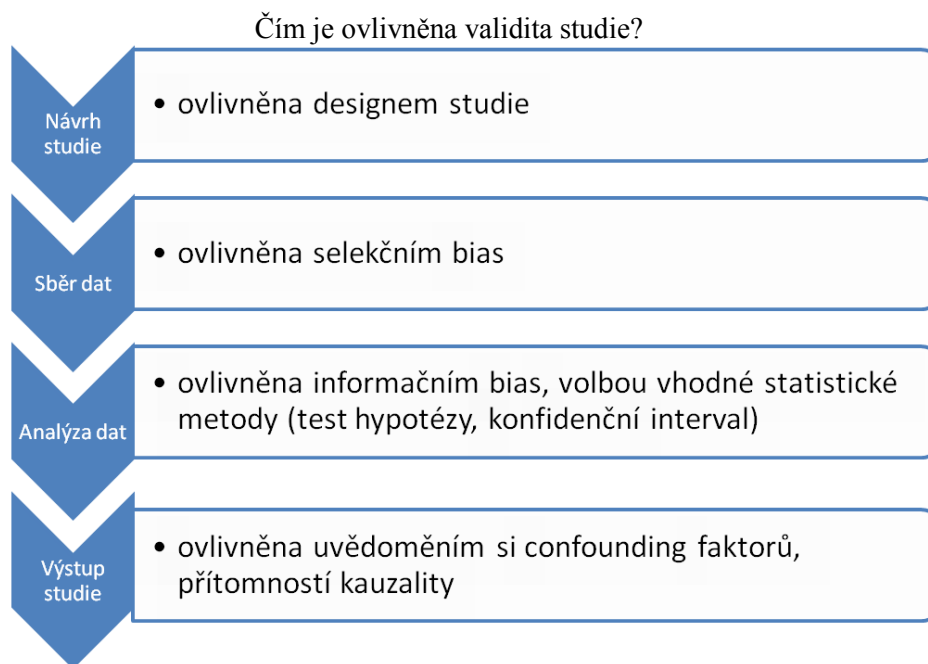


## Interpretace výsledků studií

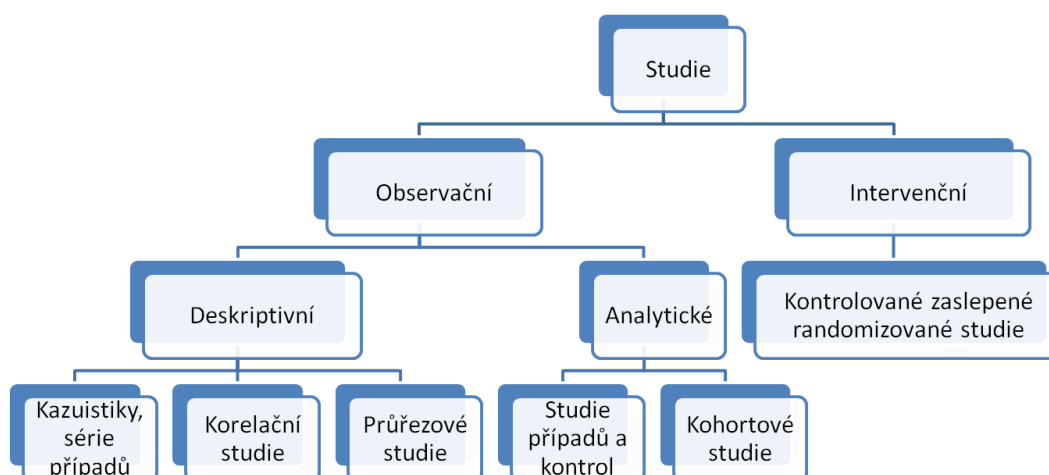
Výsledky studie je třeba interpretovat komplexně. Většina typů studií, které se zabývají kvantitativním výzkumem, předkládá určitý výstup, jež poskytuje informaci o tom, zda zkoumaný jev vyšel v souladu s předpokládanou hypotézou, nebo nikoliv. Kromě výstupu, který apriorně nabízí interpretaci závěru studie, je při ní třeba zohlednit i validitu studie, „věrohodnost“, validní je taková studie, která nemá nedostatky ve svém designu, ve sběru dat a v jejich analýze.



### 1. Design studie

Studií budeme rozumět buďto šetření (když badatel nezasahuje do průběhu zkoumání, jen pozoruje, též observační studie), nebo pokus (experiment, též intervenční studie; když badatel zasahuje do průběhu zkoumání, ovlivňuje různé parametry působící na zkoumaný jev). Příkladem šetření je pozorování četnosti karcinomu plic u kuřáků a nekuřáků (badatel se jen „dívá“ – vyhledává v dokumentaci, anamnéze pacienta informaci o kouření), příkladem pokusu je klinické hodnocení léků (badatel volí skupiny pacientů, kterým dá lék, zasahuje do jejich terapie, zkouší různé léky apod.).

Podstatné je, že už samotný design (návrh, styl) studie ovlivňuje to, jak moc lze výsledkům takové studie přikládat důležitost, validitu. Výhody a nevýhody jednotlivých studií, které je nutné zohlednit při interpretaci, ukážeme na přehledu studií.



Výhody a nevýhody:

**Kazuistiky, série případů** – pro nás nejsou příliš zajímavé, jsou čistě popisné, výhodné pro popis vzácných jevů – raritní genetické syndromy apod., kde není možné nasbírat dostatečné množství probandů pro statistickou analýzu; nevýhoda: závěr je subjektivní a čistě věcí autora (nízká validita), nebo dokonce chybí.

**Korelační studie** – porovnávají jeden znak s druhým v rámci jednoho výběru, např. míra spotřeby alkoholu v závislosti na přítomnosti kolorektálního karcinomu, výhoda: rychle vypracovány („za jedno odpoledne“), často dávají atraktivní závěry (s oblibou v bulvárních časopisech – studie typu „lidé, co jí čokoládu, žijí déle“ - taková studie vznikne jednoduše tak, že stát s velkou spotřebou čokolády (např. Švýcarsko; informace pochází třeba z tržních statistik) má hodně dlouhožijících obyvatel (informace ze zdravotních souhrnných statistik), ale skutečně ti, co jedli čokoládu, žili nejdéle? apod.); nevýhoda: velmi nízká validita, neboť obvykle ne ti samí probandi, co mají jeden znak (např. spotřebovávají alkohol), mají i druhý znak (karcinom), čili závislost tak nemůže být prokázána (např. Česká republika má světové prvenství jak v pití piva, tak v četnosti kolorektálního karcinomu, nicméně většina lidí, co pije pivo, karcinom nemá, naopak většina, co ho má, nepije pivo...).

**Průřezové studie** – jsou podobné korelačním.

**Studie případů a kontrol** – zkoumají souvislost mezi expozicí určitému faktoru a jeho následkem (např. kouřili lidé (expozice), co mají momentálně karcinom plic(následek)?); jsou retrospektivní, tedy „díváme“ se u jejich zpracování do minulosti – dokumentace apod., proto jsou výhodné u častých následků a vzácných expozic (o jejichž přítomnosti se dozvíme z dokumentace); hodnotí se pomocí odds ratio (OR), další výhody: možnost hodnocení více expozic na jeden následek, rychlé, levné; nevýhoda: často problém s selekčním, informačním bias (viz dále), nemožnost sledovat real-time rozvoj patologického mechanismu expozice-následek, nevhodné u vzácných následků.

**Kohortové studie** – zkoumají vztah mezi momentálně přítomnou a nadále působící expozicí a časem se rozvinuvším následkem, jsou tedy prospektivní (máme výběr kuřáků a nekuřáků a sledujeme, u koho se rozvine častěji karcinom plic); výhodné u vzácných expozic, možnost sledovat více následků jedné expozice a real-time sledování patologického mechanismu, klesá riziko špatné informace ze špatně vedené dokumentace (informační bias), nevýhodou časová náročnost (roky), nevhodné u vzácných následků, riziko ztráty probandů („léta trvající studie je přestane bavit“), hodnocena dle risk ratio (RR).

**Kontrolovaná zaslepená randomizovaná studie** – vrcholný typ studie, zkoumá real-time vliv expozice (rizikového, častěji projektivního) na rozvoj určitého faktoru (např. vliv léku A a léku B na zlepšení dané nemoci), vše pod dohledem badatele; kontrolovaná = přítomnost kontroly (zamezení desinterpretace z absence „normy“), zaslepená = probandi netuší, zda jsou ve skupině pod vlivem expozice, nebo ne (léčení lékem A, nebo B?), randomizovaná = na začátku všichni

probandi jsou porovnatelní, co se týče objektivních faktorů (např. všichni mají stejnou nemoc), co se týče neměřitelných faktorů (adherence k léčbě, skryté expozice, rodinná zátěž apod.), musí být náhodně rozděleni do dvou skupin, kdy neměřitelné faktory jsou rovnoměrně distribuovány do obou skupin (léčené, neléčené) a neovlivní výsledek studie.

Závěry studií je tedy nutné interpretovat kriticky (např. byť bude výsledek korelační studie sebeatraktivnější, je nutné ho vždy brát s rezervou). V pořadí, v jakém jsou jednotlivé typy studií uvedeny, se zvyšuje i jejich validita daná designem. Nejvalidnější závěry tak dává kontrolovaná randomizovaná zaslepená studie, ne všude ji však lze využít (omezení vyplývá především z toho, že se jedná o intervenční studii, a zasahovat můžeme pouze eticky – není možné vědomě a uměle vystavovat probanda rizikovému faktoru jen kvůli výsledku, zda je expozice skutečně škodlivá).

## 2. Bias

Bias (čti 'bajəs, česky zkreslení, ale nepřekládá se) je někdy nazýván systematickou chybou, jedná se jakýkoliv vliv na sběr, analýzu a interpretaci dat, který zkreslí závěr studie (obvykle sníží validitu, někdy ale paradoxně posílí závěr; viz dále). Bias se tedy uplatňuje prakticky ve všech částech studie, dle jejich fází se různě nazývá: během sběru selekční bias, během analýzy dat informační bias, během interpretace interpretační bias (častěji zvaný confounding).

Bias může závěr studie zkreslit tak, že jej podhodnotí, nebo nadhodnotí. Podhodnocení je lepší varianta, např. zjistíme-li, že výsledek „šance na rozvoj karcinomu plic u kuřáka je 2,5krát větší než u nekuřáka“ je podhodnocen, pak bude šance ve skutečnosti pro kuřáka ještě vyšší a závěr je tak prokazatelnější. Bias sám o sobě nelze měřit, lze pouze odhadnout jeho směr – zda posouvá odhad závěru směrem k nadhodnocení, nebo podhodnocení. V podstatě každá studie je do různé míry zatížena přítomností bias, jeho role by měla být vždy v závěru diskutována (ve stati Diskuze).

Nyní probereme konkrétní příklady bias a jeho vliv na závěr studie, a tedy interpretaci:

### Selekční bias:

- *Volunteer bias* = bias dobrovolníka, proband, který se do studie sám nadšeně přihlásí, bývá častěji ten s následkem, nemocí (slibuje si, že mu studie od nemoci pomůže), obvykle i lépe a více sleduje svůj zdravotní stav, spíše si tak vzpomene na přítomnost expozice než provane bez následku, který sice také mohl být exponován (např. ozářením v minulosti), ale nekladl na to zřetel => vliv expozice může být v závěru nadhodnocen
- *Recall bias* = bias tázaného, proband si snáze vzpomene na skutečnosti, které ho před tazatelem nekompromitují, resp. takové skutečnosti uvede v pravé míře (např. pacient záměrně podhodnotí svou konzumaci alkoholu ve studii o vlivu alkoholu na rozvoj kolorektálního karcinomu, zatají svou sexuální promiskuitu při studii o vlivu sexuální aktivity na šíření venerické infekce apod., zde je vliv expozice podhodnocen; jindy zase proband záměrně nadhodnotí svou míru sportovní aktivity, např. při studii o vlivu sportování na kardiovaskulární nemoci apod., expozice je pak nadhodnocena); řešením je přesně definovat kritéria pro expozici, následek, využívat dokumentace a výpovědi okolí před výpovědí probanda
- *Interviewer bias* = bias tazatele, tazatel obvykle (byť podvědomě) stráví více času u probanda s následkem (nemocí) než u probanda bez následku, tím pádem u něho častěji zjistí přítomnost expozice, ta je tak nadhodnocena, řešením je vyškolení a nezájatost tazatelů, příp. využití normovaných dotazníků při sběru dat

### Informační bias

Uplatňuje se při analýze dat, kdy je nutné správně odlišit probandy s expozicí a následkem, s expozicí bez následku, bez expozice s následkem a bez expozice a bez následku. Pro to je nutné přesně desinovat všechny stavy: s expozicí/bez ní a s následkem/bez něho. Někdy jsou sem řazeny i zkreslení spadající pod selekční bias.

Pokud probandi bez následku stejně jako s následkem nadhodnotí/podhodnotí expozici, expozice je tak zkreslena v obou skupinách a výsledek bude podhodnocen (lepší varianta), neboť bude zkreslení „rovnoměrně“ rozloženo v obou skupinách (tzv. nediferenciální chyba);

pokud ale nadhodnotí/podhodnotí expozici jen jedna ze skupin (buďto s následkem, nebo bez něj), závěr bude nadhodnocen (horší varianta; tzv. diferenciální chyba).

Riziko, které může ovlivnit výsledek, je ztráta probandů (a tedy i informací), ať už z důvodu jejich úmrtí, zkoumáme-li např. výskyt akutních kardiovaskulárních příhod, nebo z důvodu ztráty adherence (též compliance, probandova zájmu) probanda o studii (v informovaném souhlasu je vždy uvedena možnost, že proband může ze studie kdykoliv odstoupit). Přesáhne-li ztráta probandů cca 30% výchozího množství, studie ztrácí smysl a není jiného východiska, než ji předčasně ukončit. Naštěstí je možné použít metodu nelepšího a nejhoršího scénáře k tomu, abychom obdrželi určité výsledky – např. při kohortové studii vlivu pití kávy na incidenci akutního infarktu myokardu, která bude trvat deset let, přestane studie určité procento probandů bavit, pokud bude toto procento rovnoměrně rozloženo mezi skupiny pijící i nepijící kávu, diferenciální chyba není přítomna, pokud ale přestane studie atrahovat jen ty, co kávu nepijí, je expozice kávě neúměrně nadhodnocena a objeví se diferenciální chyba. Tu je možné odstínit metodou nejhoršího a nejlepšího scénáře – nejhorší scénář znamená, když budeme uvažovat, že všichni, co kávu nepili a odpadli, by měli kardiovaskulární příhodu, nejlepší znamená, že by ji všichni zmínění neměli. Pak je výsledkem nikoliv jedna konkrétní hodnota relativního rizika pití kávy vs. kardiovaskulární příhody, ale interval, kde nižší číslo odpovídá výpočtu z nejhoršího scénáře (káva, která způsobí kardiovaskulární příhody, je podhodnocena všemi příhodami, které uměle přiřkneme všem těm, kteří kávu nepili a odpadli), a vyšší číslo náleží výpočtu z nejlepšího scénáře (všem, co kávu nepili a odpadli, zamítneme možnost, aby utrpěli kardiovaskulární příhodu, káva a její vztah k příhodám ve skupině pijáků je tak nadhodnocen). Skutečná hodnota relativního rizika leží někde mezi oběma čísly, tedy uvnitř intervalu.

**Interpretační bias (confounding)** – viz dále.

### 3. Výstup studie (zvolená statistická metoda – hladina významnosti, konfidenční interval, poměr šancí a relativní riziko)

*Test hypotézy* = matematický algoritmus, který ve výsledku vrací pravděpodobnost pracovní (alternativní) hypotézy proti nulové hypotéze. Alternativní hypotézu obvykle chceme dokázat (resp. lépe řečeno chceme ukázat, že jejím přijetím se nedopouštíme výrazné chyby). Nulovou hypotézou může být např. „incidence karcinomu plic je u kuřáků i nekuřáků shodná“, zatímco alternativní hypotézou pak „incidence karcinomu plic je u kuřáků jiná (konkrétněji vyšší) než u nekuřáků“. Výsledek testu hypotézy je pravděpodobnost chyby, kterou činíme, pokud přijmeme za platnou alternativní hypotézu. Této pravděpodobnosti se říká hladina významnosti. Pro naše účely je možné test hypotézy chápat jako určitou „černou skříňku“, která je nutnou součástí statistické analýzy, ale jako uživatele (čtenáře studie) nás zajímá jen její výstup – hladina významnosti.

*Hladina významnosti* = kritérium určující statisticky významný rozdíl zkoumané veličiny mezi dvěma; porovnávanými soubory, příp. závislost dvou znaků jednoho výběru, např. efekt léku A u prvního souboru astmatiků a léku B u druhého souboru astmatiků, příp. závislost váhy na výšce pacientů; značí se  $p$  (též  $p$ -value,  $p$ -level). Interpretace: „ $p$  = pravděpodobnost chyby prvního typu, přijmeme-li za platný závěr, že se oba soubory signifikantně liší“. Konvenčně přijato:  $p \leq 0,05 \Rightarrow$  zkoumaný znak se mezi oběma soubory signifikantně liší;  $p > 0,05 \Rightarrow$  zkoumaný znak se mezi oběma soubory signifikantně neliší. Neříká nic o kauzalitě závislosti.

*Interval spolehlivosti* = interval spolehlivosti (confidence interval, CI) je interval, ve kterém leží určitá skutečná hodnota (obvykle střední hodnota míry polohy) s určitou pravděpodobností, obvykle 95 % (někdy 99 %). CI lze použít v podstatě pro odhad každé veličiny (průměr, OR, RR), např. je-li 95% CI RR roven (1,25; 2,50), pak skutečné RR (relativní riziko) leží mezi hodnotami 1,25 a 2,50 s pravděpodobností 95%. Je-li rozložení veličiny  $X$  normální, pak lze 95% CI sestavit z odhadnuté (spočítané) střední hodnoty veličiny, z odhadnuté odchylky  $s$  a rozsahu výběru  $n$  ( $X$  může být průměr, OR, RR):

$$95\% \text{ CI: } X \in \left( \hat{X} - 1,96 \cdot \frac{s}{\sqrt{n}}; \hat{X} + 1,96 \cdot \frac{s}{\sqrt{n}} \right)$$

- konfidenční interval (CI) se zvětšuje:
  - s rostoucí variabilitou ( $\uparrow s$ ) – tu moc neovlivníme
  - s rostoucí pravděpodobností ( $\uparrow$  kvantil) – 95% kvantil = 1,96; 99% kvantil = 2,58 atd. (<http://new.euromise.org/czech/tajne/ucebnice/html/html/node15.html>)
  - s klesajícím rozsahem souboru ( $\downarrow n$ )
- a naopak CI se zmenšuje s  $\uparrow n$ ,  $\downarrow s$ ,  $\downarrow$  pravděpodobnost.

*Poměr šancí a relativní riziko*

- poměr šancí (odds ratio, OR)
  - užívá se u studií případů a kontrol (case-control study)
  - poměr šancí následku v exponované a neexponované skupině

$$OR = \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{ad}{bc}$$

- relativní riziko (relative risk, RR)
  - užívá se u kohortových a intervenčních studií
  - poměr incidencí následku v exponované a neexponované skupině

$$RR = \frac{\frac{a}{a+b}}{\frac{c}{c+d}} = \frac{a(c+d)}{c(a+b)}$$

	efekt ano	efekt ne	
expozice ano	<i>a</i>	<i>b</i>	<i>a+b</i>
expozice ne	<i>c</i>	<i>d</i>	<i>c+d</i>
	<i>a+c</i>	<i>b+d</i>	<i>a+b+c+d</i>

- je-li OR nebo RR:
  - =1 => expozice nemá na výskyt efektu žádný vliv (pouze náhodný)
  - >1 => expozice zvyšuje pravděpodobnost efektu; např. expozice tabákovému kouří zvyšuje riziko/incidenci karcinomu plic
    - typické pro rizikové expozice
  - <1 (ale >0) => expozice snižuje pravděpodobnost efektu, např. expozice slunečnímu záření snižuje riziko/incidenci křivice
    - typické pro protektivní expozice (faktory).

#### 4. Confounding (interpretační bias)

Confounding (čti kon'faundin, (z)matení; opět se nepřekládá) značí přítomnost dalších faktorů, které se značnou měrou (někdy i kauzální!) podílejí na rozvoji následku. Expozice tak není jediným faktorem, nebo není vůbec faktorem, který na rozvoji následku participuje. Uvedme příklad: Máme dva výběry, v jednom z nich probandy, kteří nepijí vůbec mléko (vyhledáme je jako pacienti trpící deficitem laktázy), v druhém probandy, kteří mléko pijí. Sledujme, v kterém z nich se častěji rozvine karcinom plic (jedná se o kohortovou studii). Po deseti letech zjistíme, že mezi probandy, kteří pili pravidelně mléko, se objevil karcinom plic signifikantně častěji. Je nyní možné říct, že mléko je rizikový faktor v rozvoji karcinomu plic?!

Na první pohled se nám zdá zjištěná závislost zarážející, neodpovídá naší zkušenosti. Skutečně je zde dobře nechat se vést intuicí, neboť odpověď na otázku zní „ne, na základě provedené studie nelze považovat mléko za rizikový faktor v rozvoji rakoviny plic“. Proč? Kromě mléka totiž mohlo na probandy, kteří pravidelně kravský produkt konzumovali, působit i mnoho dalších faktorů, např. kouření. Kdybychom se každého ze zúčastněných zeptali, zda kouří, pravděpodobně

bychom zjistili, že mezi pijáky mléka je mnohem více kuřáků než mezi těmi, co mléko nepili. Skutečným, kauzálním faktorem, který vedl k rozvoji rakoviny, tak bylo kouření, nikoliv mléko. Pití mléka zde hraje roli „matoucího faktoru“, confounderu.

Takových faktorů může hrát v každé studii roli celá řada, účinnou obranou je buďto stratifikace, tedy rozdělení výběrů dle jednotlivých expozičních faktorů na ty, kteří jsou, a ty, kteří nejsou exponováni, nevýhodou je pak to, že vznikne mnoho malých skupinek, kde se obtížně prokazuje signifikace. V našem případě (kohortové studii) by bylo správné stratifikovat obě skupiny (ti, co pili, i nepili mléko) na kuřáky a nekuřáky (vzniknou tak čtyři nové menší skupiny) a porovnat pití/nepití mléka u kuřáků a nekuřáků. Pak pravděpodobně nebude mezi porovnávanými skupinami, co do pití mléka, u kuřáků i nekuřáků žádný rozdíl, což je průkaz, že mléko nehraje kauzální roli v rozvoji karcinomu plic a jen „mate“ (je confounderem). Stratifikaci je možné použít především u observačních studií.

Druhou možností, použitelnou pouze u kontrolovaných zaslepených randomizovaných studií, ale o to účinnější, je randomizace = náhodné rozdělení probandů na vstupu do studie do dvou skupin, čímž se různě přítomné confoundery jednotlivých probandů rovnoměrně distribuují do obou skupin (z nichž např. jedné budeme podávat lék a druhé placebo), čili jejich výsledný vliv na výsledek se bude navzájem rušit. To je možné jen u intervenční studie, neboť náhodné rozhození probandů do dvou skupin by u observačních studií vedlo i k chaotickému rozdělení probandů, které by jistě nerespektovalo rozdělení přítomnosti/nepřítomnosti expozice, a studie by tak nemohla zkoumat vliv expozice (ta by byla rovnoměrně rozdělená do obou skupin, což nechceme, my chceme jednoznačné dělení probandů do dvou skupin právě podle přítomnosti/nepřítomnosti expozice).

Třetí možností je matematické modelování, které se může pokusit odstínit vliv confounderů, nicméně jedná se o náročnou a specifickou problematiku.

## 5. Kauzalita

Kritéria kauzality stanovil v roce 1965 A. B. Hill a jedná se o soubor pravidel, která by měla být mezi expozicí a následkem, resp. mezi dvěma porovnávanými znaky splněna, abychom mohli postulovat jejich příčinnou závislost. Statistická potvrzená závislost (především dostatečně nízká hladina významnosti) totiž ještě neznamená, že mezi porovnávanými znaky existuje i závislost kauzální. Statistická analýza pouze odhaduje roli náhody, která může být zavržena i tehdy, pokud není kauzalita přítomna (viz příklad s mlékem a karcinomem plic, kde jsme rádobu signifikaci prokázali, ale kauzalita byla na první pohled vyloučena). Méně nápadným případem může být, pokud prokazujeme signifikantní rozdíl mezi dvěma výběry, které jsou dostatečně velké, pak je relativně snadné statisticky významný rozdíl prokázat, byť nemusí být reálný, resp. praktický rozdíl přítomen. Máme-li tedy určitou závislost mezi dvěma znaky (i mezi expozicí a následkem) pokládat za příčinnou, kauzální (rozumějme „opravdovou“), pak je nutné kromě signifikantní statistické analýzy subjektivně zhodnotit a potvrdit přítomnost následujících kritérií kauzality (nejlépe všech):

- *Síla asociace* = silnější asociace (souvislosti mezi dvěma znaky, expozicí a následkem) jsou pravděpodobnější.
- *Konzistence* = požaduje soulad výsledku naší studie s výsledky studií opakovaných za stejných podmínek jinde.
- *Specifičnost* = jedna příčina (expozice) má vést vždy k jednomu a tomu samému následku.
- *Časový sled (temporalita)* = příčina (expozice) má vždy přecházet následkem. Jedná se o jediné kritérium, které je podmínkou nutnou k ustanovení kauzální závislosti; ostatní kritéria nemusí být potvrzena, a přesto se může jednat příčinný vztah.
- *Plausibilita* = biologická přijatelnost, pochopitelnost, přiměřenost
- *Experimentální důkaz* = jeho výsledek musí odpovídat výsledku studie. Mnohdy ale není z reálných nebo etických důvodů proveditelný.
- *Koherence* = soulad výsledku s dosavadním věděním o problematice, dosavadními závěry, které již prošly kritérii kauzality a byly přijaty.
- *Analogie* = asociace (zkoumaná závislost) je pravděpodobnější tehdy, je-li podobná jiné asociaci (závislosti).