

EPIDEMIOLOGICKÉ STUDIE II.

Mgr. Aleš Peřina, Ph. D.

První část tématu věnovaného epidemiologickým studiím (odučeno 5. 3. 2020) bylo věnováno typům dat, které se používají pro účely epidemiologického výzkumu a také tomu, jak lze data získávat. Druhá část se věnuje možnostem zpracování dat.

ODDÍL 1: Strategie deskriptivní statistiky

■ Spojitá proměnná

- *Ordinální: když 2 je více než 1: lze vyjádřit na stupnici, aritmetické operace mají smysl (můžeme sečíst výšku všech osob ve studii)*
- *Intervalová: rozpětí ordinálních hodnot: typicky např. indexy např. BMI)*

■ Nespojité (kategoriální) data

- *Nominální stupnice (somatotyp, krevní skupina, barva očí...), nelze provádět aritmetické operace (např. sčítat dvě krevné skupiny, ale můžeme vyjádřit počet osob, které mají daný znak, např. krevní skupinu A.*

Deskripce spojité proměnné

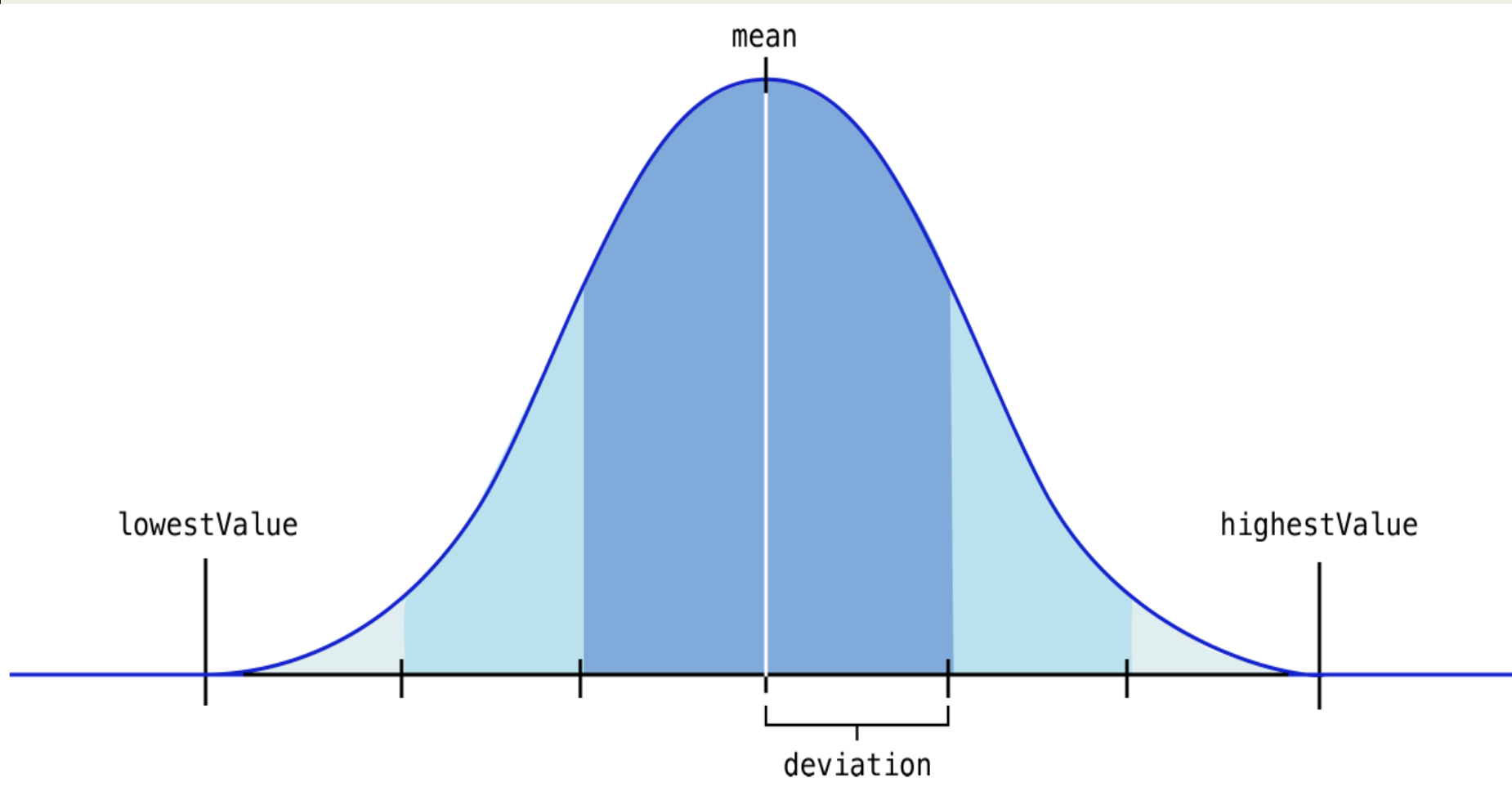
- Průměr: součet množiny čísel vydělený počtem těchto čísel, ukazatel centrální tendence
- Medián: prostřední hodnota, 50. percentil v řadě čísel seřazených vzestupn
- Modus: nejčastěji se vyskytující hodnota v soubor
- Minimum: nejmenší hodnota v soubor
- Maximum: největší hodnota v souboru

Většina výsledků je produktem jednoduchých aritmetických operací s čísly.

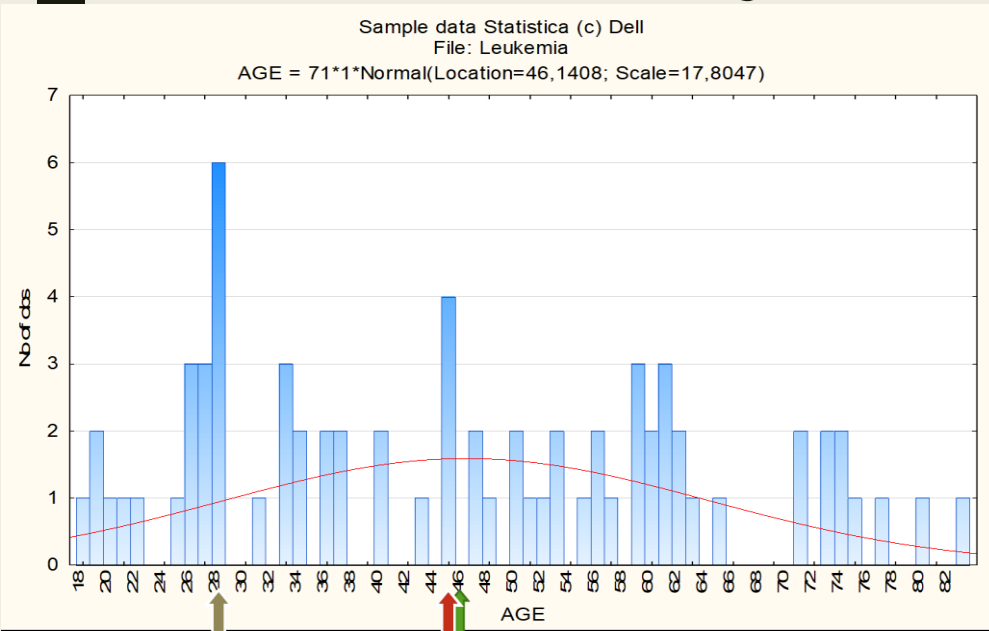
U symetrického rozložení platí: průměr = medián = modus

Průměr

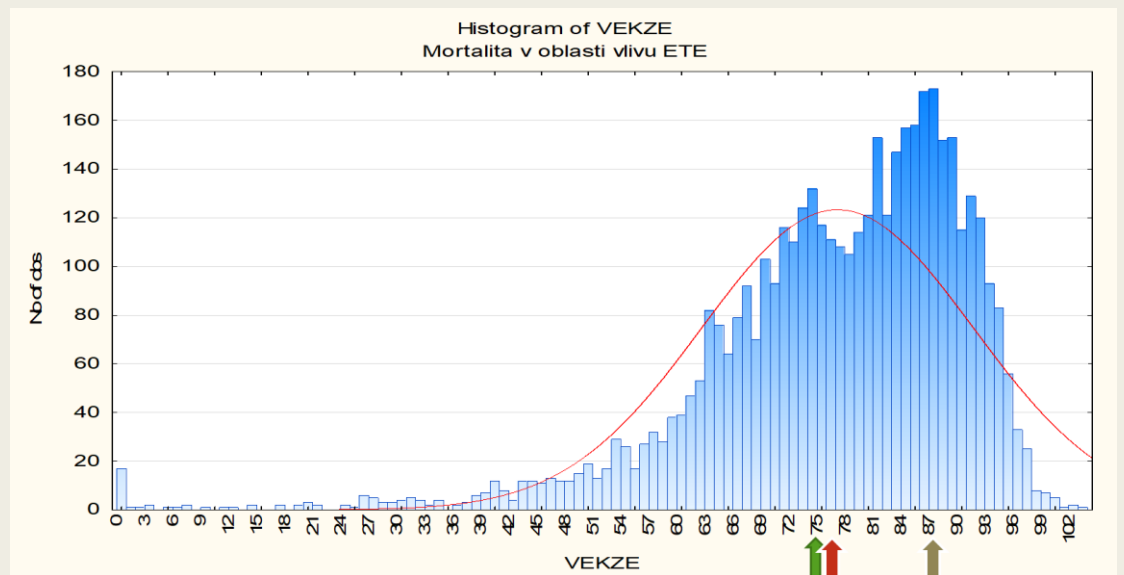
Použití průměru má však svoje pravidla: je-li soubor dostatečně velký a jev dostatečně náhodný, pak má tendenci seskupovat se kolem průměru: většina lidí má normální hmotnost.



Úskalí použití průměru na nesourodých datech



Na reálných datových sadách vidíme řadu nepravidelností, viz graf vlevo. Na souboru vidíme, že neexistuje nějaká typická hodnota, kolem které by se ostatní shlukovaly, takže by bylo použití průměru nevýmluvné. V těchto případech je proto věrohodnější použití tzv. neparametrických ukazatelů, tj. mediánu, modu, minimálních hodnota a maximálních hodnot.



Deskripce nespojitých dat

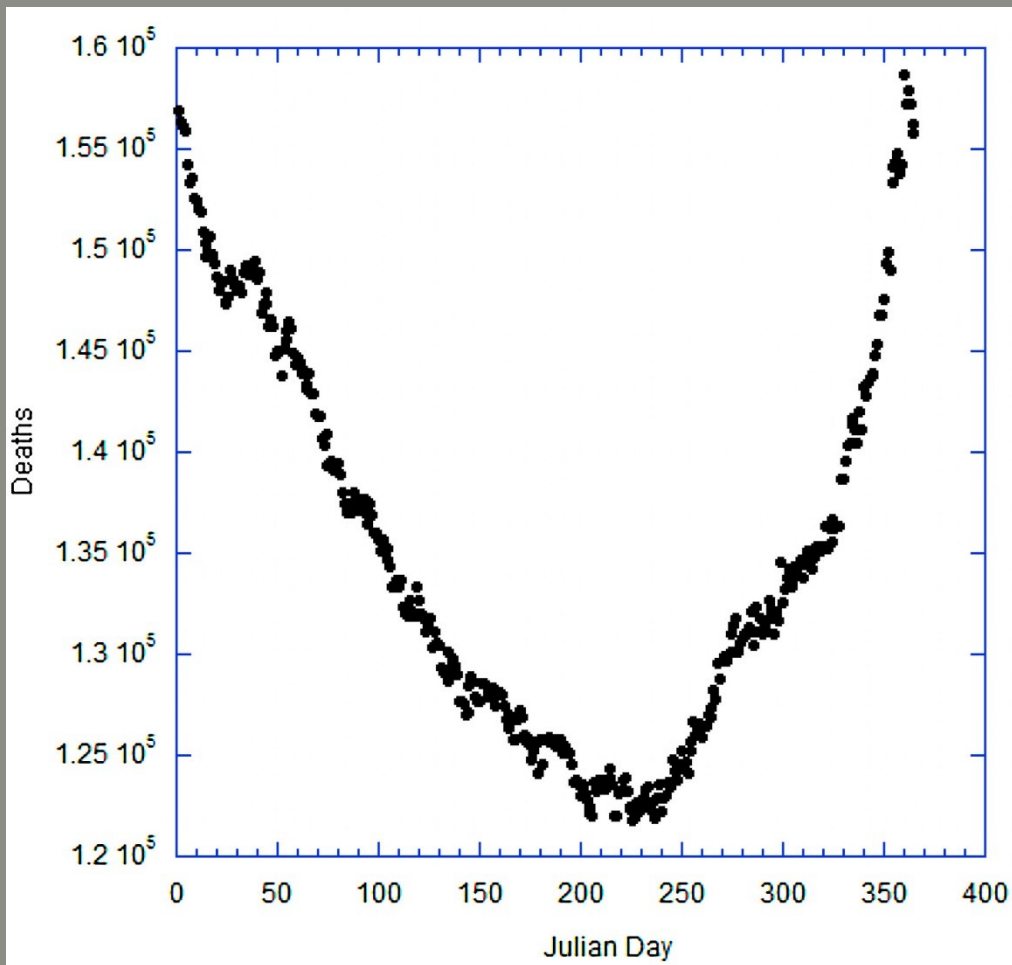
Data nelze sčítat, násobit či průměrovat, popisujeme vyjádřením četností (počtu případů, které vyhovují definovanému znaku).

- Četnost: počet výskytu hodnoty dané kategorie
- Kumulativní četnost: součet hodnot sledované kategorie a všech nižších kategorií
- Relativní četnost: počet výskytu hodnot vztažených ke zvolenému základu

Příklad frekvenční tabulky

Konzumujete mléčné výroby?	Absolutní četnost	Kumulativní četnost	Relativní četnosti
ano	52	52	56,6 %
Jen zakysané	20	$52 + 20 = \mathbf{72}$	21,7 %
Jen z kozího mléka	14	$52 + 20 + 14 = \mathbf{86}$	15,2 %
ne	6	$52 + 20 + 14 + 6 = \mathbf{92}$	6,5 %
CELKEM	92	92	100,0 %

Vyjádření relativních četností umožňuje porovnávat nestejně velké soubory (celek je vždy 100 %).



Sezónní úmrtnost obyvatelstva v USA (Grant WB et al., 2017)

I relativně malý rozdíl na ose Y může mít v určitých souvislostech vypovídající hodnotu.

Ukázka výtěžnosti deskriptivní studie

Ukázka výtěžnosti jednoduché deskriptivní studie

Jedná se o shromážděné údaje o počtu zemřelých v USA v závislosti na době, která již uplynula od začátku roku. Lze vidět, že nejvíce zemřelých je na začátku a na konci roku, nejméně zemřelých pak uprostřed roku v letních měsících.

Jev se vysvětluje výskytem respiračních onemocnění v zimních měsících (všimněte si malého „zubu“ kolem 50. dne v roce na ose x; chřipková epidemie). Naopak v letních měsících může být lepší dostupnost např. čerstvých potravin s obsahem mnoha ochranných faktorů.

Podobné výkyvy v úmrtnosti lze pozorovat i v České republice.

Expozice a následek

■ Expoziční proměnná

- *Spotřeba potravin, frekvenční dotazník, spotřeba vody, fyzická aktivita.na*
- *O zdraví přímo nevyovídá, ale zpravidla na základě literární rešerše můžeme usuzovat, že zdraví ovlivňují pozitivním způsobem*

■ Znak: následek expozice

- *Zdraví je definováno WHO jako stav úplné tělesné, duševní a sociální pohody, nejen nepřítomnost nemoci nebo vady.*
- *Nemoc: zkoumání vyžaduje přesnou definici diagnostického kritéria, jinak hrozí zkreslení studie (tzv. inclusion a exclusion kritéria)*
- *Smrt*

Zdraví, nemoc a smrt vyjadřujeme ukazateli, jako nemocnost a úmrtnost, vztažená na jednotku populace.

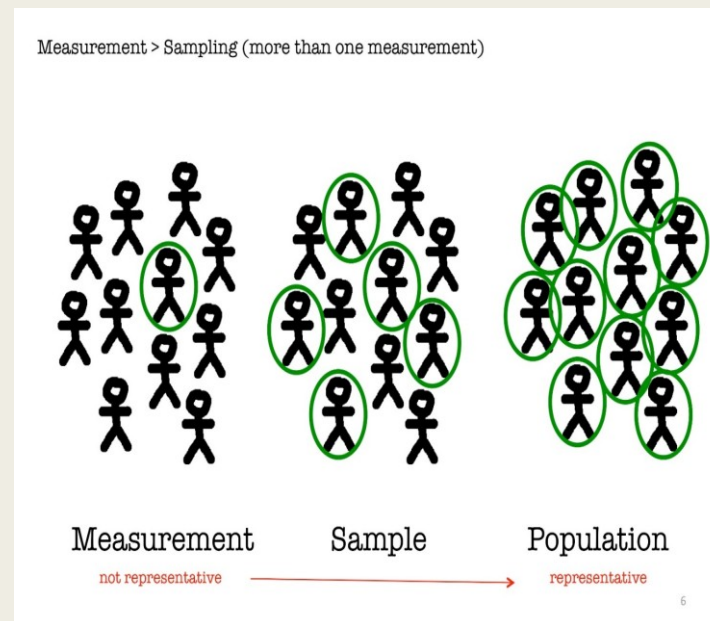
Incidence a prevalence: způsob vyjadřování následku

- Incidence: nově vzniklé případy
 - *Osobočas je zvláštním ukazatelem využívaným zejména v populačních studiích na bázi rutinně sbíraných dat. Vyjadřuje počet účastníků studie násobený délkou studie*
 - 30 osob během 2-leté kohortové studie je 60 osobooboroků
- Prevalence: existující případy k určitému datu (bodová) nebo období (intervalová; např. během roku).
 - *Prevalenci snižují vyléčení nebo zemřelí.*
 - *Prevalenci zvyšují chronicky nemocní.*
- Incidence a prevalence se vyjadřují v relativních číslech vztažených k vhodnému základu (obvykle 100 tis. osob)

Studovaná populace

Nejpřesnější údaje lze získat vyšetřením každé osoby z populace. Hovoříme o tzv. vyčerpávajícím šetření. To však bývá neúměrně nákladné a mnohdy neproveditelné. Provádíme proto:

- **reprezentativní výběr** tak, aby reprezentativní vzorek populace svoji strukturou napodoboval základní populaci.
- Kritéria pro dosažení reprezentativnosti:
 - *Randomizace (nahodilost) výběru. Každý musí mít stejnou šanci stát se prvkem výběrového souboru.*
 - *Stratifikace (rozvrstvení) tak, aby procentuální zastoupení různých vrstev populace (tzn. pohlaví, podíl osob se základním, středním a vysokým vzděláním atd.) bylo podobné, jako v základní populaci.*



Senzitivita a specificita

Studované znaky zdraví musí být validní, senzitivní a specifické:

■ Validita

- *Popisuje studovaný znak, situaci, kterou hodláme studovat? Ptáme se:*
 - Lze z frekvence konzumace potravin skutečně usuzovat na zdravotní stav? Není třeba zúžit skupinu sledovaných potravin?

■ Senzitivita a specificita

- *Senzitivita: úspěšnost záchytu sledovaného znaku*
 - Test na okultní krvácení (přítomnost krve ve stolici) je velmi senzitivní, protože zachytí i velmi malé množství krve
- *Specificita: schopnost testu odlišit případy, které znak nemají*
 - Test na okultní krvácení je velmi málo specifický, protože jeho účelem je rozpoznat nemocné s kolorektálním karcinomem, avšak velmi mnoho diagnostikovaných kolorektálním karcinomem netrpí.

Senzitivita a specificita

Vztah mezi senzitivitou a specificitou lze vyjádřit i graficky. Zapamatujte si zejména, co vyjadřují jednotlivé výroky na průniku řádků a sloupců.

	Má pozitivní výsledek testu	Má negativní výsledek testu
Je nemocný/-á	Skutečně pozitivní (SENZITIVITA)	Falešně negativní
Je zdravý/-á	Falešně pozitivní	Skutečně negativní (SPECIFICITA)

Typy studií

- Deskriptivní studie: popisují výskyt studovaného jevu, vhodné pro formulaci hypotéz.
- Analytická studie pracují se zkoumaným a kontrolním souborem, tyto se vzájemně porovnávají:
 - *Ekologická: měří korelaci mezi dvěma znaky, např. spotřebu potravin a incidenci nemocí, vycházejí zpravidla z populačních dat o úmrtnosti a nemocnosti.*
 - *Případů a kontrol (case – control) je pravidla retrospektivní. Na základě tzv. inclusion a exclusion kritérií určíme osoby s přítomností znaku a nemoci. Kontrolním souborem jsou pak osoby, které znak (nemoc) nemají. Pátráme po rozdílech v **expozičních proměnných**.*
 - *Kohortová je zpravidla prospektivní a dobré kohortové studie trvají řadu let. Na začátku studie jsou definovány populační skupiny, u nich se přivoředpokládají odlišné expoziční znaky (studenti, ekonomicky aktivní lidé, chroničtí nemocní). V průběhu studie se sledují měnící se životní a pracovní podmínky a způsob života, Nakonec zhodnotíme, do jaké míry různé expozice ovlivnily zdraví.*
- Kontrolovaný experiment
 - *Na populační úrovni prakticky nerealizovatelný. Nejčastěji se využívá v klinickém výzkumu. Jedna skupina osob dostává lék, druhá neúčinné placebo a porovnáváme dopad na zdraví (vyléčení nemoci).*
 - *Kontrolovaný experiment podléhá náročným etickým kritériím.*

Analytické studie, smyšlený příklad:

Výsledky studie shromáždíme do čtyřpolní tabulky, jak ukazuje horní tabulka. První možností vyhodnocení je výpočet pozorovaných a očekávaných četností, jak znázorňuje dolní tabulka. Zvyšující se rozdíl mezi pozorovanými a očekávanými četnostmi může být důkazem existence závislosti. Ve smyšleném příkladu je rozdíl pouze 12-ti osob na souboru 509 pacientů. Tento postup je nezávislý na typu použité studie.

Pozorované četnosti	Zdroje Fe nedostatečné	Zdroje Fe dostatečné	CELKEM
Anemie ANO	205	129	334
Anemie NE	89	86	175
CELKEM	294	215	509

Očekávané četnosti	Zdroje Fe nedostatečné	Zdroje Fe dostatečné	CELKEM
Anemie ANO	193 = $294 \times 334 / 509$ - 12	141 + 12	334
Anemie NE	101 + 12	74 - 12	175
CELKEM	294	215	509

Měření velikosti efektu pomocí ukazatelů velikosti rizika

Tentýž příklad můžeme vyhodnotit také pomocí ukazatelů rizika jednoduchým vyjádřením poměrů (postup výpočtu je naznačen pod tabulkou):

Pozorované četnosti	Zdroje Fe nedostatečné	Zdroje Fe dostatečné	CELKEM
Anemie ANO	205	129	334
Anemie NE	89	86	175
CELKEM	294	215	509

Poměr šancí (Odds Ratio, u studií případů a kontrol)

1. krok: $205/89 = 2,30$

2. Krok $120/86 = 1,50$

Výsledek

$OR = 2,30/1,50 = 1,53$

O 53 % větší šance.

Relativní riziko (u kohortových studií)

1. krok: $205/294 = 0,70$

2. krok $125/215 = 0,60$

Výsledek:

$RR = 0,70/0,60 = 1,17$

1,17-krát nebo o 17 %

více

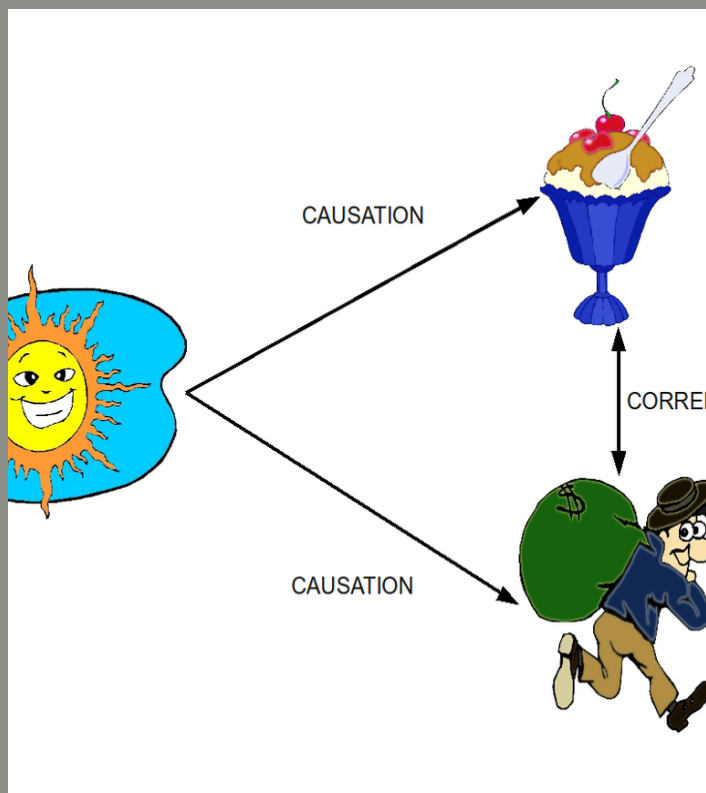
Obecný formát 2x2 (čtyřpolní) tabulky

Exponovaní Nemocní	Neexponovaní Nemocní <i>Falešně pozitivní</i>
Exponovaní Zdraví <i>Falešně negativní</i>	Neexponovaní Zdraví

RR nebo OR = 1 ... není efekt
RR nebo OR > 1 ... rizikový faktor, poškozují zdraví
RR nebo OR < 1 ... ochranný faktor, podporuje zdraví

Zapamatujte si význam číselky 1 při vyjadřování rizika. Když riziko vypočítáme jako poměr sloupců tabulky, pak pouze poměr totožných čísel je roven 1. Totožné výsledky znamenají neexistenci rozdílu mezi výzkumnými soubory.

Bias a confounders

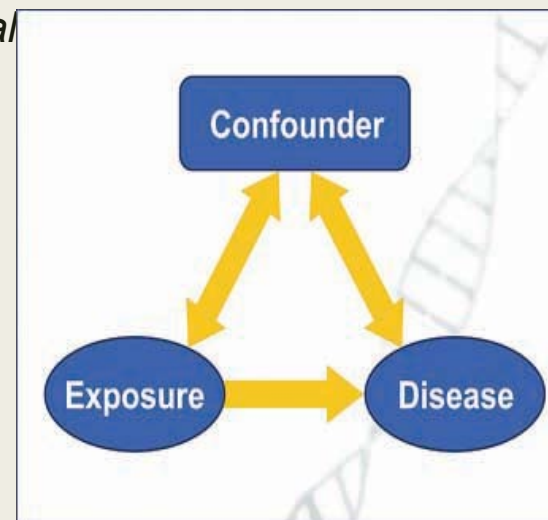


Prosté matematické vyjádření výsledků musí být interpretováno s velkou pečlivostí a opatrností (náplň diskusí vědeckých prací).

Chyby bývají způsobeny vzájemnými interferencemi mezi expozičními znaky a studovanými znaky, které jsou následkem expozičních faktorů.

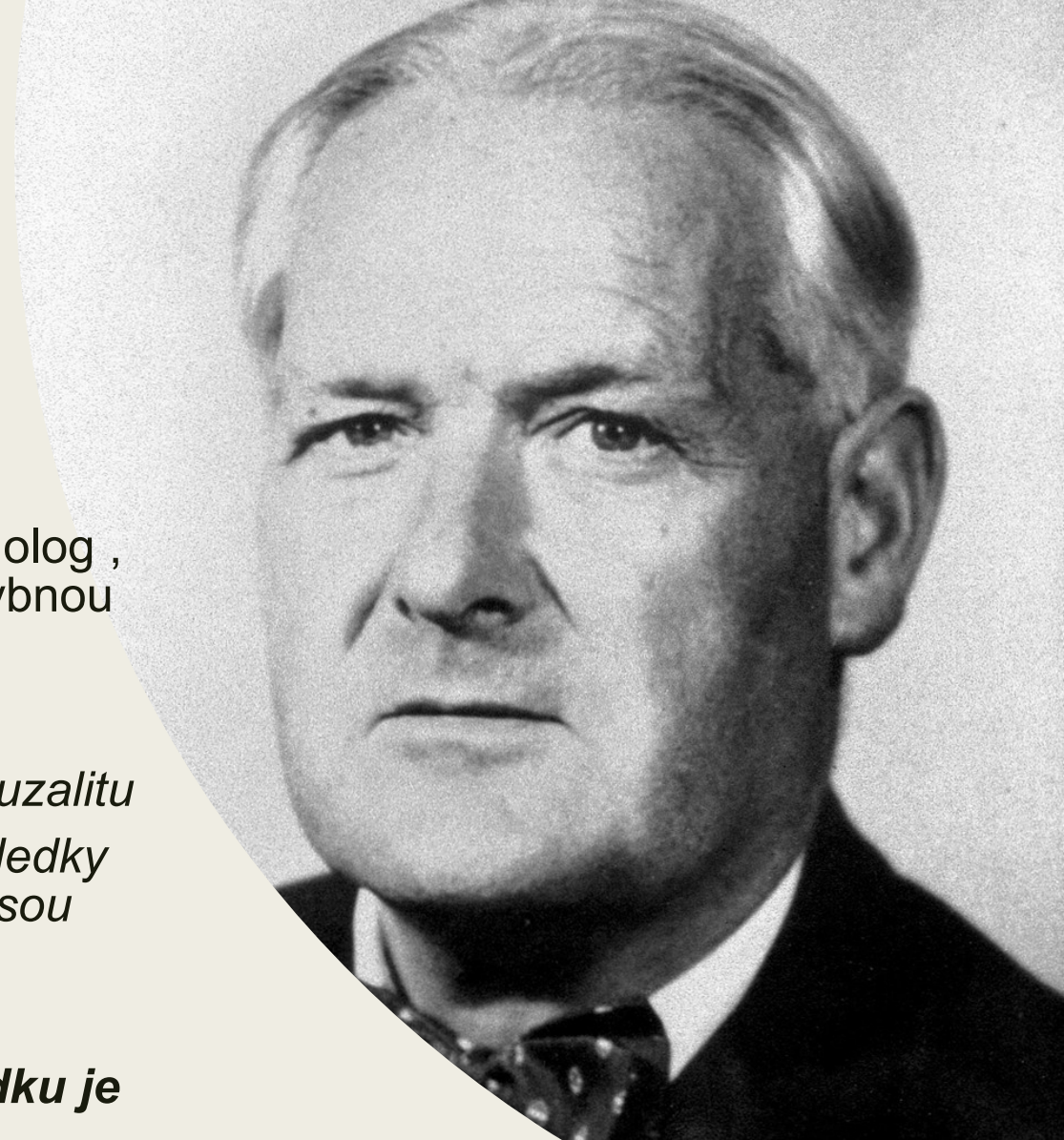
Důsledkem je pak vznik chyby, chybné interpretace výsledku studie. Jev mohl být způsoben jiným faktorem, který jsme nerozpoznali.

Spotřeba zmrzliny matematicky koreluje s nárůstem kriminality ve městě, ale ve skutečnosti je nárůst kriminality způsoben teplem v letních sezónách.



Epidemiologie v hodnocení zdravotních rizik

- sir Austin Bradford Hill, (1897 – 1991), britský epidemiolog , zabýval jsem studiem chyb, které mohou způsobit chybnou interpretaci výsledků epidemiologických studií.
- Hillova kritéria kauzality
 - **Síla asociace:** ani slabá asociace nevylučuje kauzalitu
 - **Konzistence:** kauzalitu může podpořit, když výsledky různých epidemiologických studií na totéž téma jsou obdobné
 - **Specificita** není podmínkou
 - **Přísná časová posloupnost expozice a následku je základní platnou podmínkou!**



Epidemiologie v hodnocení zdravotních rizik II.

■ Hillova kritéria kauzality ... pokrač.

- **Biologický gradient:** obvykle silnější působení expozičního znaku znamená i silnější znak sám (větší konzumace vlákniny snižuje riziko kolorektálního karcinomu), nemusí však platit za všech okolností.
- **Biologická přijatelnost:** schopnost patofyziologického vysvětlení je žádoucí.
- **Koherence:** kompatibilita se "zavedenými" teoriemi je výhodná, není-li však nalezena, můžeme stát na prahu nového objevu
- **Experimentální důkaz je žádoucí.**
- **Analogie:** její absence může být jen projevem nedostatku vědecké představitosti

