

# C7188 Úvod do molekulární medicíny 6/12



## MOLEKULÁRNÍ EPIDEMIOLOGIE



**Ondřej Slabý, Ph.D.**

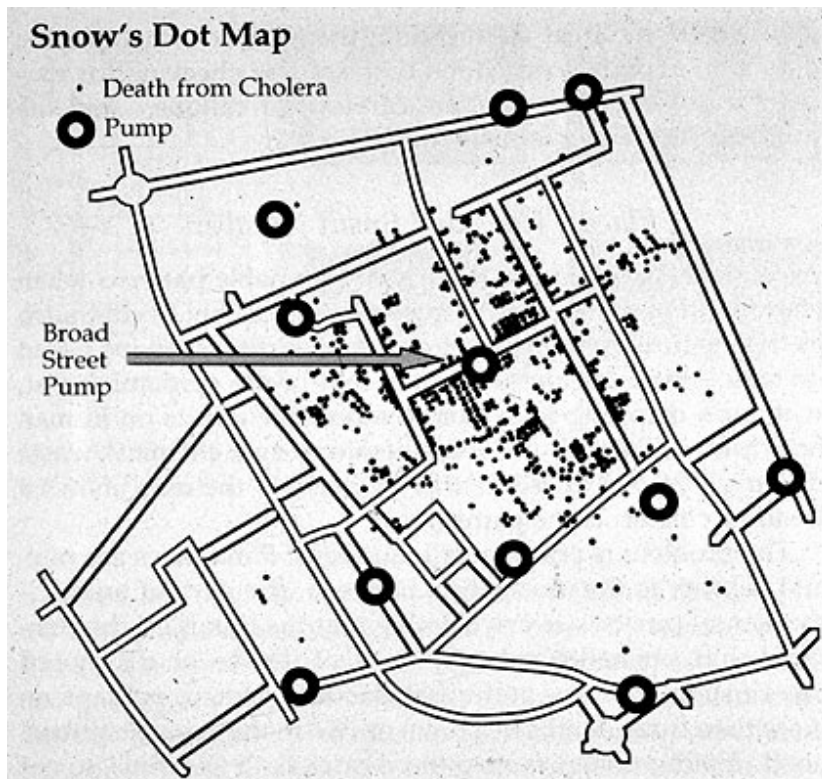
*Masarykův onkologický ústav*

*Univerzitní centrum buněčné imunoterapie*

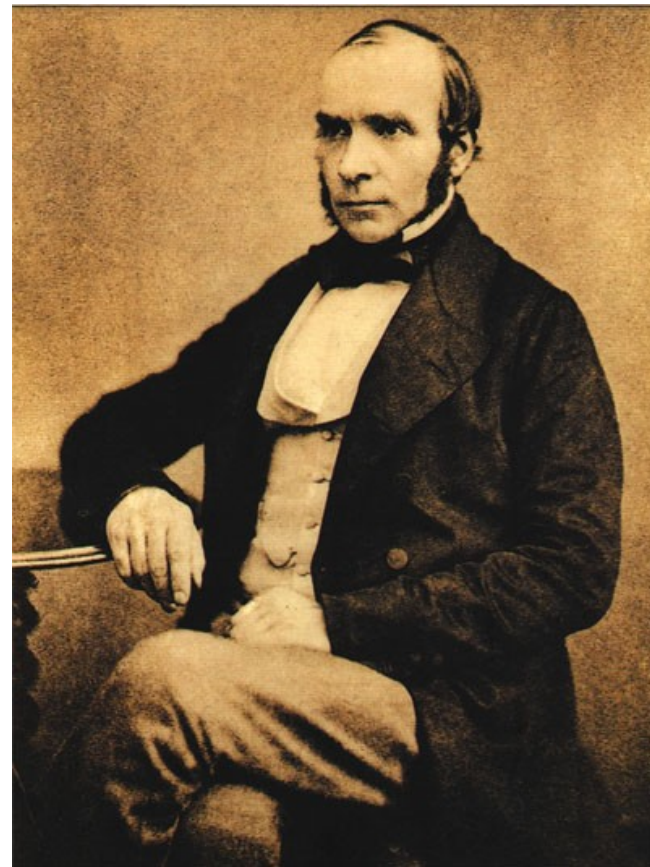
*Lékařská fakulta Masarykovy univerzity*



# Historie – cholera v Londýně (1854)



Of 83 people, only 10 lived closer to a different pump than Broad Street  
Of these 10, 5 preferred taste of Broad Street water and 3 were children who went to nearby school



John Snow is credited by many with developing the modern field of epidemiology

# John Snow and the Pump Handle

John Snow and cholera  
in 1854 London

<http://www.ph.ucla.edu/epi/snow.html>



# Ignác Filip Semmelweis (1818 -1865)

Horečka omladnic v porodnické klinice ve Vídni  
vznik choroby přičítán  
„kosmicko-telurickým" vlivům

**Zavedl mytí rukou v chlorové vodě**

Příčiny, pojem a ochrana před horečkou omladnic

**„Vražďení šestinedělek musí přestat!„**



# Epidemiologie – vymezení pojmu

Původ slova epidemiologie je odvozen z řeckých slov *epi* (nad, mezi) a *démos* (lid) a *logos* (slovo, věda, studium) a může být volně přeložen jako „studium toho, co je nad lidmi“.

- **popis zdravotního stavu populace, popis frekvence a rozsahu a onemocnění na populační úrovni**
- **popis vlivu onemocnění na populaci, včetně socioekonomických dopadů**
- **identifikace příčin a rizikových faktorů určitého onemocnění, poskytuje podklady pro preventivní opatření**
- **hodnocení efektivity lékařské péče, jak preventivní tak i následné, poskytuje podklady pro management**
- **prognózy vývoje onemocnění**
- **podklady pro analýzu rizik a stanovení zákonných limitů**

# Epidemiologie – vymezení pojmu

## ➔ Incidence

- popisuje výskyt nových případů onemocnění v populaci ve studovaném období
- počet nových případů / celkový počet osob ve studované populaci v určitém časovém úseku
- vysoké hodnoty incidence ukazují na vysoké riziko onemocnění
- výhodné při popisu akutních onemocnění

## • Prevalence

- popisuje zastoupení dané nemoci ve studované populaci v daném okamžiku
- okamžitý počet nemocných / celkový počet osob ve studované populaci
- prevalence výhodná pro popis dlouho trvajících chronických onemocnění
- prevalence závisí na počtu osob, které onemocněli a na délce nemoci
- není nutně mírou rizika onemocnění

$$\text{Prevalence} = \text{Incidence} \times \text{Doba trvání}$$

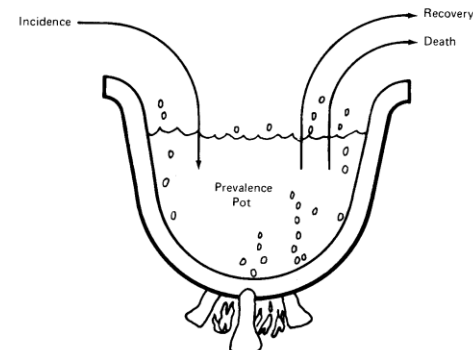


FIGURE 5 The Relationship between Incidence and Prevalence.

*Přednáška částečně převzata od Ing. Miloslava Pouzara Ph.D.*

# Kauzální vztah

- ➔ Pokud změna frekvence, nebo kvality expozice vede k odpovídající změně ve frekvenci výskytu onemocnění
  
- ➔ Typy kauzálních vztahů
  - postačující příčina – pokud je daný faktor přítomen, nemoc se vždy projeví (genetická onemocnění – Downův syndrom)
  - nutná příčina – pokud daný faktor není přítomen, nemoc se neprojeví (infekční nemoci – tuberkulóza)
  - rizikový faktor – pokud je daný faktor přítomen, zvyšuje pravděpodobnost vzniku onemocnění (cigaretový kouř – rakovina plic)
  - přímý kauzální vztah vs. nepřímý kauzální vztah
  - nekauzální vztah – mezi proměnnými je náhodná (nevysvětlující) závislost (lineární vztah mezi počtem zubních plomb a rizikem infarktu myokardu)

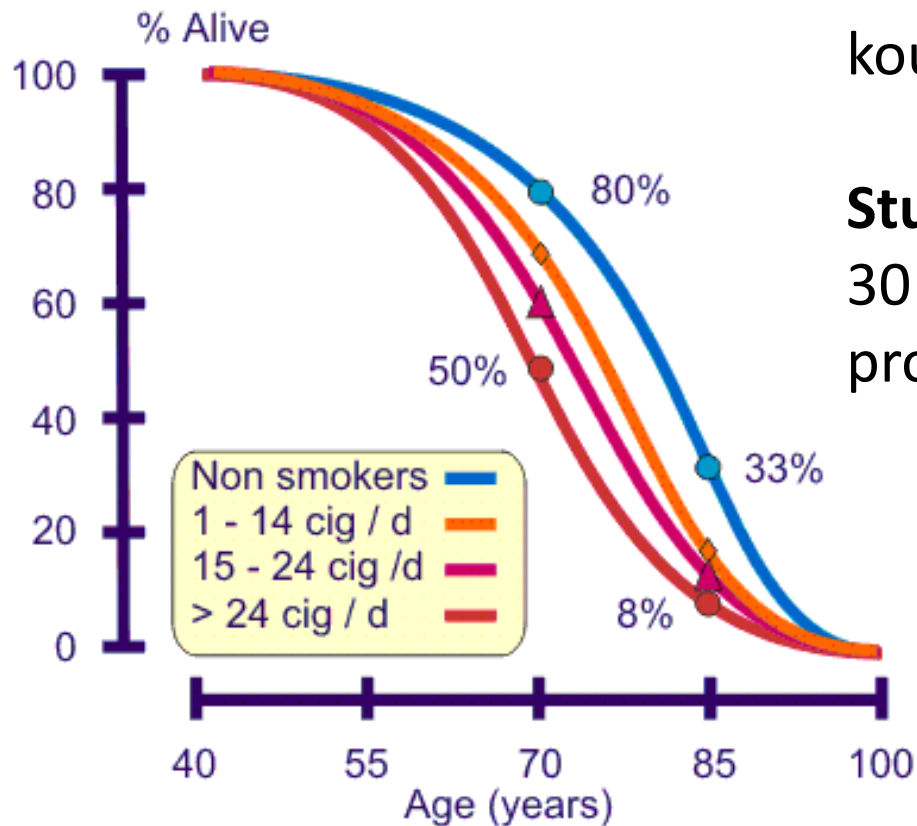
## Austin Bradford Hill (1897-1991)

Jako první prokázal vztah  
kouření a karcinomu plic

### Studie na britských lékařích

30 000 účastníků

prospektivní 50 let trvající studie

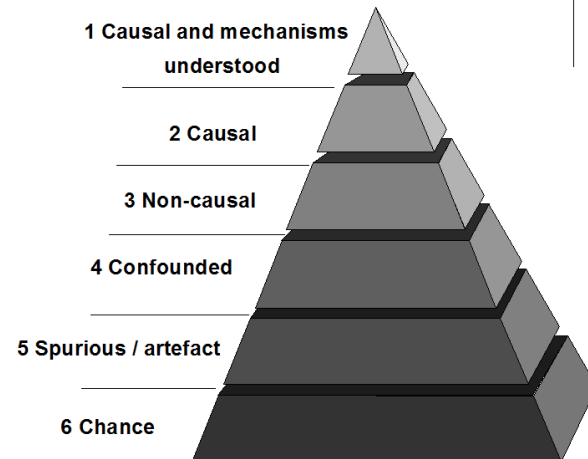




# Bradford Hillova kriteria kauzality

- **Korelace (síla asociace)** – čím větší korelace mezi nezávislou proměnnou (expozice) a závislou proměnnou (počet onemocnění), tím vyšší pravděpodobnost kauzálního vztahu mezi proměnnými
  - Semelweis (1818-1865) - významně vyšší úmrtnost na horečku omladnic na klinice řízené lékaři, oproti porodnici s porodními bábami – příčinná souvislost?
  - Durkheim 1951 – počet sebevražd ve 4 pruských regionech na poč. 19 století koreluje s poměrným zastoupením protestantského obyvatelstva – příčinná souvislost ?
- **Časová souslednost** – příčina by měla předcházet následek
  - blesk předchází hrom – příčinná souvislost?
  - v letech kdy se začalo s průmyslovým spalováním odpadů vzrostl počet případů rakoviny – příčinná souvislost? (doba latence 20 let !)
  - v letech po zákazu používání azbestu stále stoupá počet případů azbestem vyvolané rakoviny – příčinná souvislost? (doba latence 20 let)

Figure 5.13 The pyramid of associations



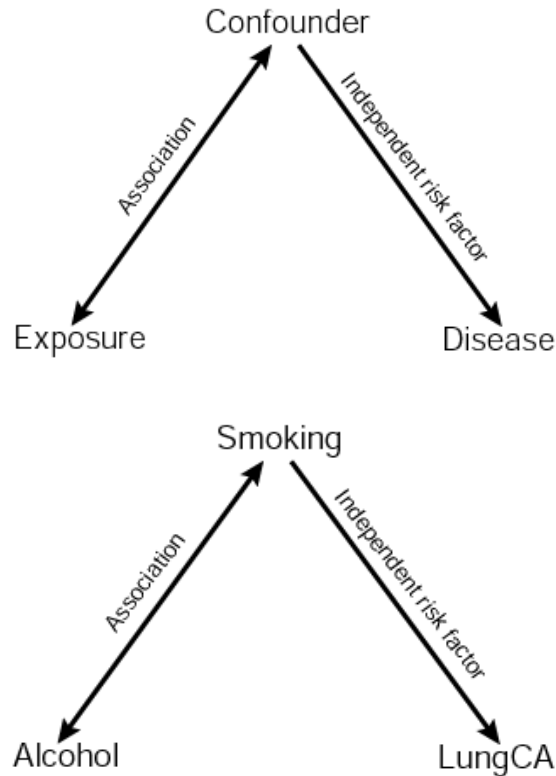
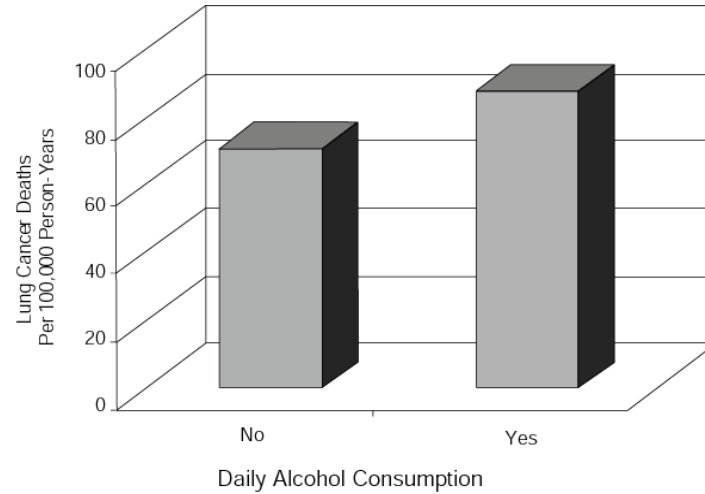
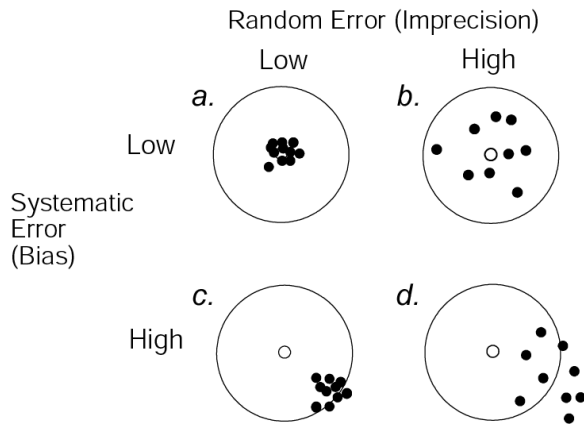
# Bradford Hillova kriteria kauzality

- **Konzistence** – čím větší je shoda výsledků různých studií zabývajících se danou dvojicí
- **Koherence** – logická provázanost s výsledky jiných vědních oborů
- **Biologická přijatelnost (věrohodnost)** – existence teoretického mechanismu vysvětlujícího vztah mezi příčinou a následkem
  - formaldehyd je genotoxický a vysoce dráždivý – informace o zvýšeném výskytu tumorů v dýchacích cestách je věrohodná
  - formaldehyd má velký distribuční objem a rychle se v organismu odbourává, ani vysoké koncentrace v ovzduší výrazně nezvyšují jeho koncentraci v tělních tekutinách – informace o zvýšeném výskytu nádorů vnitřních orgánů nevěrohodná
- ⇒ **Specifická** – čím menší počet příčin postačuje k vysvětlení následku, tím lépe
- ⇒ **Vztah dávka účinek** – čím větší míra expozice tím větší míra a četnost následků
  - emise z diesellových motorů – mnoho studií popisuje vztah k rakovině plic, není vztah dávka účinek, patrně vliv kouření (confounding factor)

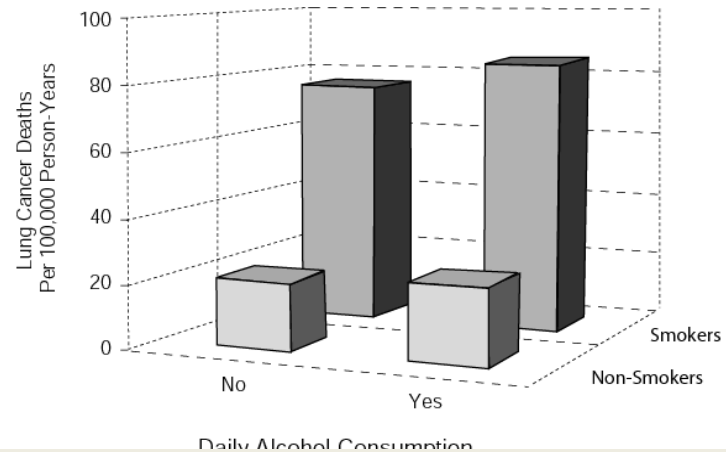
# Faktory ovlivňující určení kauzality

<b>Bias</b> (zkreslení – systematická chyba)	Vychýlení výsledku jedním směrem (zvyšuje pravděpodobnost nalezení neexistujícího vztahu) - Selection bias (výběrové zkreslení), Allocation bias (chyba přiřazení), Recall bias (informační zkreslení), publikační bias, observační bias, bias daný ztrátou osob ze studie, apod.
<b>Random Error</b> (náhodná chyba)	snižuje pravděpodobnost nalezení existujícího vztahu, snižuje sílu statistických testů
<b>Confounding factor</b> (zavádějící faktor)	neexistující vztah mezi expozicí látky A a onemocněním je nalezen díky vztahu mezi expozicí látky A a látky B, která skutečně dané onemocnění vyvolává
<b>Synergismus</b>	vzájemná kombinace dvou či více příčin a jednoho následku
<b>Modifikace efektu</b>	různé úrovně expozice vyvolávají různá onemocnění

**Rakovina plic o 30% častější mezi konzumenty alkoholu**



**Stratifikace dat – zvlášť kuřáci a nekuřáci**



**kouření je zavádějící faktor při popisu vztahu mezi konzumací alkoholu a rakovinou plic**

# Epidemiologické studie

## Observační (pozorovací)

**Kauzistiky** (*case reports*)

**Série případů** (*case series*)

**Korelační studie** (*corelational studies*)

**Průřezové studie** (*cross-sectional studies*)

DESKRIPTINÍ

**Studie případů a kontrol** (*case-control studies*)

**Kohortové studie** (*cohort studies*)

ANALYTICKÉ

## Intervenční (experimentální)

**Klinické studie** (*clinical trial*)

**Terénní kontrolované studie** (*field trial*)

K  
V  
A  
L  
I  
T  
A  
  
V  
Ý  
P  
O  
V  
Ě  
D  
I

# Průřezové studie

## Cross-sectional studies

- **retrospektivní studie**
- náhodný výběr jedinců ze studované populace – data na individuální bázi
- jednorázový sběr dat o zdravotním stavu a expozici studovaných jedinců
- prevalence onemocnění a prevalence expozic (**varianty daného genu**)
- rozdělení populace na 4 skupiny – exponovaní s nemocí, exponovaní bez nemoci, neexponovaní s nemocí a neexponovaní bez nemoci
- v jednom časovém bodě porovnáváme riziko onemocnění v exponované a neexponované populaci nebo míru expozice v nemocné a zdravé populaci
- můžeme najít vztah (asociaci), ale nelze prokázat kauzalitu
- nedostatek informací o časové souslednosti, možná záměna příčiny a následku (různá skladba potravin u lidí s určitou nemocí – příčina? následek?)
- retrospektivní sledování – zkreslení informací zejména o expozici
- rychlost, nízká cena

# Míry efektu – průřezová studie

**Risk disease** (riziko onemocnění) (exposure) –  $A/A+B$

**Risk disease** (riziko onemocnění)(no exposure) –  $C/C+D$

**Risk ratio** (poměr rizik) –  $(A/A+B)/(C/C+D)$

**TABLE 21.3 Prevalence Data: Rate Ratio and Rate Difference**

		Disease	
		+	-
Exposure	+	<i>A</i>	<i>B</i>
	-	<i>C</i>	<i>D</i>

Alelicka varianta daného genu

Prevalence of disease in exposed population =  $A/(A + B)$

Prevalence of disease in the unexposed population =  $C/(C + D)$

Rate ratio =  $A/(A + B)$  divided by  $C/(C + D)$

Rate difference =  $A/(A + B) - C/(C + D)$

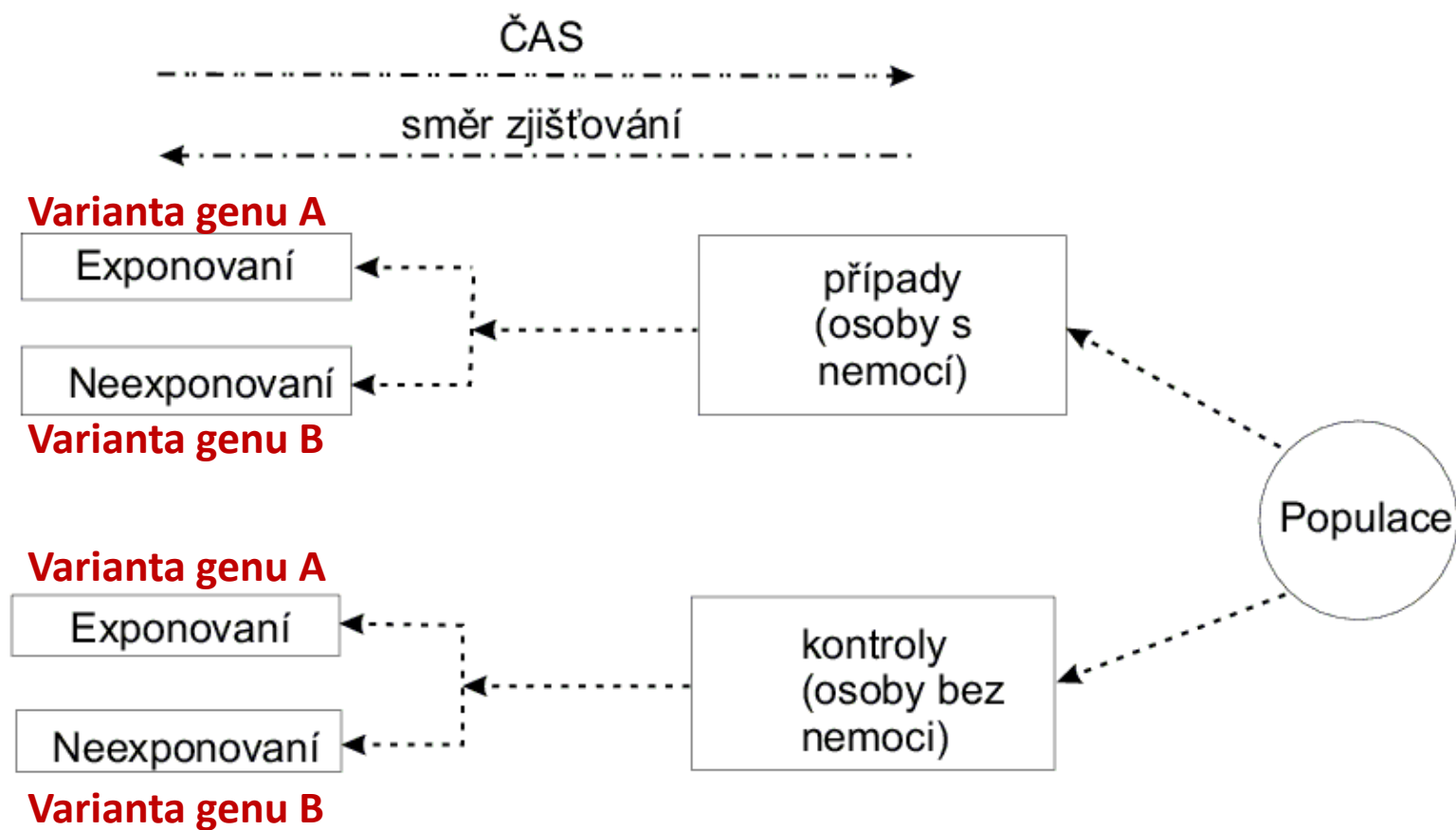
# Studie případů a kontrol

## • Case-control studies

- **retrospektivní analytická studie**
- testování hypotéz o příčinách onemocnění, o vztahu mezi expozicí a nemocností
- incidence i prevalence onemocnění, dvě populace
- **VĚTŠINA ASOCIAČNÍCH STUDIÍ (ALELA-ONEMOCNĚNÍ) MÁ CHARAKTER CASE-CONTROL STUDIE**
  
- **identifikace případů (cases)** – jedinci se sledovaným onemocněním
- **výběr kontrol (controls)** – jedinci bez sledovaného onemocnění, ale v ostatních aspektech (věk, pohlaví, socioekonomický status) co nejvíce podobní případům
  - „Density sampling“ – výběr kontrol v okamžiku identifikace případu (výhodné pro dlouhotrvající studie)
  - „Cumulative sampling“ – výběr kontrol na konci studie
  
- v obou skupinách hodnotíme úroveň expozice
  - vyšší ve skupině případů – látka působí jako rizikový faktor
  - vyšší ve skupině kontrol – látka působí jako protektivní (ochranný) faktor
  
- zkreslení
  - výběrové – skupina kontrol nereprezentuje studovanou populaci
  - „recall bias“ - případy mají tendenci nadhodnocovat expozici



## Studie případů a kontrol (case-control studies)



# Míry efektu – studie případů a kontrol

**odd disease** (šance onemocnět) (exposure) –  $A/B$

**odd disease** (šance onemocnět) (no exposure) –  $C/D$

**odds ratio** (poměr šancí) –  $(A/B)/(C/D)$

**TABLE 21.5 Case–Control Data: Odds Ratio**

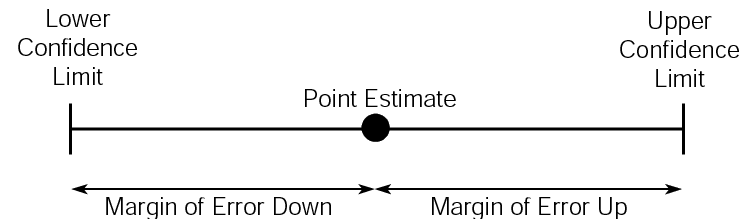
		Disease	
		+	–
Exposure	+	<i>A</i>	<i>B</i>
	–	<i>C</i>	<i>D</i>

Rate ratio = odds ratio =  $AD/BC$

## konfidenční interval (CI - confidence interval)

Interval vypočítaný na základě výběru, ve kterém je neznámý populační parametr obsažen s pravděpodobností 95%.

šířka CI – míra přesnosti (čím užší CI, tím vyšší přesnost) studie s větším počtem pozorování mívají užší CI  
 Alela A (OR, **6.3**; 95% CI, 3.4-11.6)



## Příklad studie případ-kontrola

### Stanovení OR vyjadřujícího asociaci mezi užíváním salycilátů a Reyova syndromu

	Rey. sy ano	Rey. sy ne	Celkem
Salyc. Ano	26	53	79
Salyc. Ne	1	87	88
celkem	27	140	167

**KONTINGENČNÍ TABULKA**  
Tabulka křížových klasifikací  
Čtyřpolní (2x2)

$$OR = AD/BC = 26 \times 87 / 53 \times 1 = 42,7$$

Děti s Reyovým syndromem 42,7 častěji užívaly k léčbě virového onemocnění salyciláty.

# Kohortové studie

- Cohort studies

***Kohorta je skupina osob, které mají nějakou společnou charakteristiku, a které jsou sledovány po určité období***

–V kohortových studiích jsou porovnávány skupiny zdravých osob (kohorty), rozlišené na základě přítomnosti (skupina základní) či nepřítomnosti (skupina kontrolní) suspektního rizikového faktoru.

–Prospektivní kohortové studie

- Specific exposure cohorts – kohorty tvořeny na základě různé expozice rizikového faktoru, výhodné pokud je expozice málo častá (např. Hirošima)
- General population cohorts – expoziční status jedinců stanoven až v průběhu první fáze studie, někdy periodicky přehodnocován, výhodné u častých nebo simultánních expozic

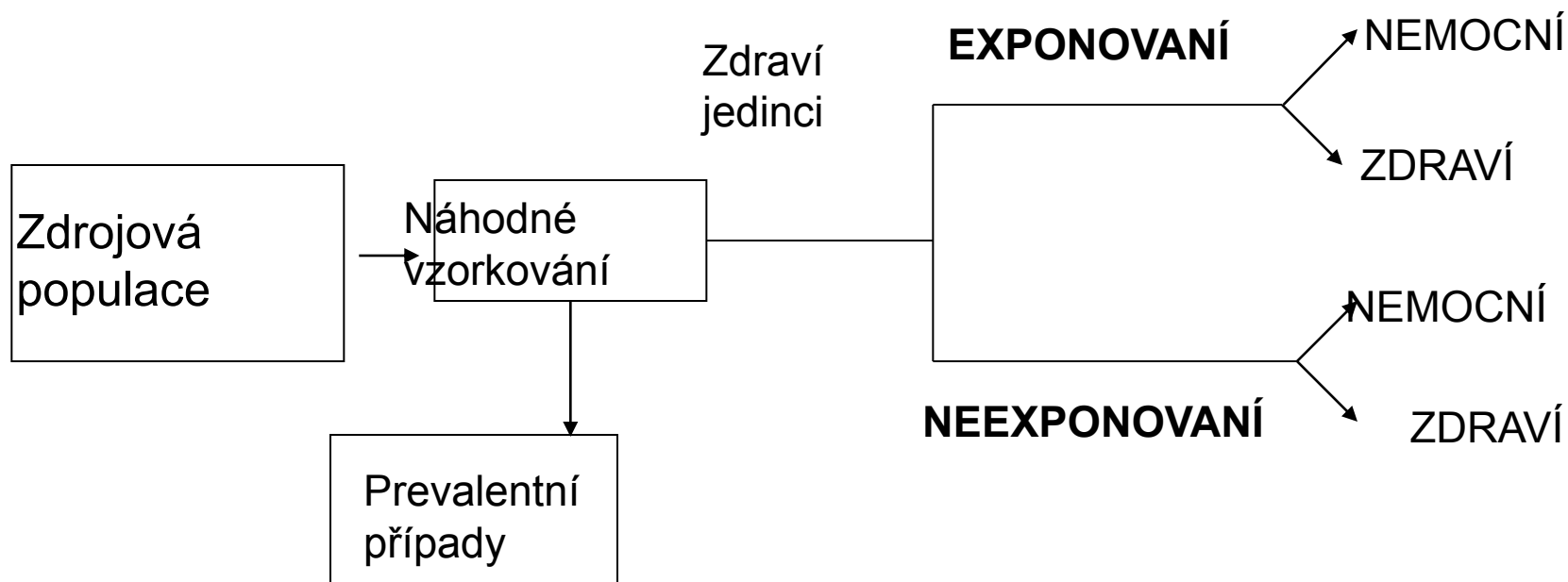
–Retrospektivní (historické) kohortové studie

–v kohortách se porovnává výskyt onemocnění

## Kohortové studie (cohort studies)

ČAS

SMĚR ZJIŠŤOVÁNÍ



# Míry efektu – kohortní studie

**TABLE 21.4 Incidence Data: Rate Ratio and Rate Difference**

		Disease	
		+	–
Exposure	+	<i>A</i>	<i>B</i>
	–	<i>C</i>	<i>D</i>

Incidence of disease in exposed population =  $A/PY$  (person-year)

Incidence of disease in unexposed population =  $C/PY$

Rate ratio =  $A/PY$  divided by  $C/PY$

Rate difference =  $(A/PY - C/PY)$

## Příklad kohortní studie

### Stanovení RR vyjadřujícího asociaci mezi kouřením matek v těhotenství a kojeneckou úmrtností

	úmrtí ano	úmrtí ne	Celkem
Kouření ano	60	1940	2000
Koření ne	80	3920	4000
celkem	140	5860	6000

Incidence exponovaných =  $60/2000 = 0,030$

Incidence neexponovaných =  $80/4000 = 0,020$

RR = poměr incidencí exponovaných/neexponovaných =  $0,030/0,020 = 1,5$

Kojenecká úmrtnost je 1,5x vyšší u dětí exponovaných.

# Síla vztahu (asociace), statistické testování hypotéz

## RR (Rate Ratio)

1.0-1.2

1.2-1.5

1.5-3.0

3.0-10.00

&gt;10.0

## Síla vztahu

žádný

slabý

střední

silný

velmi silný

- $H_0$  - neexistuje vztah mezi expozicí a onemocněním
- $H_A$  – existuje vztah mezi expozicí a onemocněním

- příklady nulových hypotéz  $H_0$

- Rate difference = 0
- RR (relative risk) = 1
- OR (odds ratio) = 1
- SMR = 100

## Závěr testu

platí  $H_0$   
(není vztah)

platí  $H_1$   
(je vztah)

nezamítnout  $H_0$ zamítnout  $H_0$ 

	platí $H_0$ (není vztah)	platí $H_1$ (je vztah)
nezamítnout $H_0$	správné rozhodnutí	Chyba 2. druhu ( $\beta$ )
zamítnout $H_0$	Chyba 1. druhu ( $\alpha$ )	správné rozhodnutí



# Definice molekulární epidemiologie

**Molekulární epidemiologie spojuje vysoce citlivé laboratorní techniky vyvinuté v rámci molekulární biologie s klasickými epidemiologickými postupy.<sup>1</sup>**

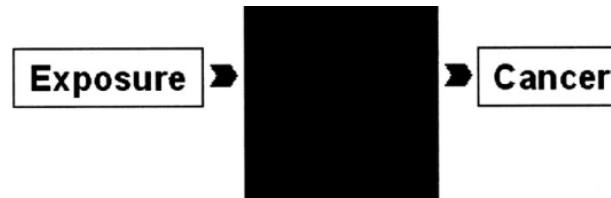
Molekulární epidemiologie se vyvíjela jako spojení mezi základním výzkumem v molekulární biologii a epidemiologickými studiemi zaměřenými především na vznik nádorového bujení u lidí. Jejím smyslem je kombinace laboratorních stanovení interní dávky, biologicky účinné dávky, biologických účinků a vlivu vnímavosti jednotlivce s epidemiologickými metodami.

Tento přístup představuje přirozenou konvergenci mezi molekulární biologií a epidemiologií.

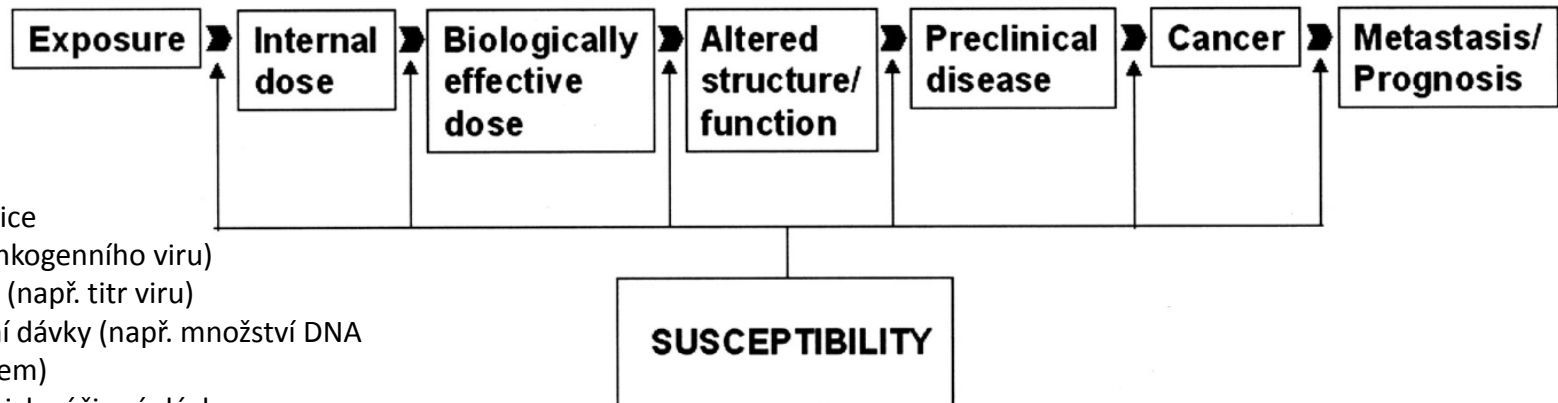
Pubmed – 2009 – 7700 odkazů v databázi při zadání „molecular epidemiology“  
Především molekulární typizace infekčních agens

<sup>1</sup>Schulte PA. *Molecular Epidemiology: Principles and Practices*. San Diego: Academic Press; 1993.

# Molekulární epidemiologie nádorových onemocnění



“TRADITIONAL EPIDEMIOLOGY”

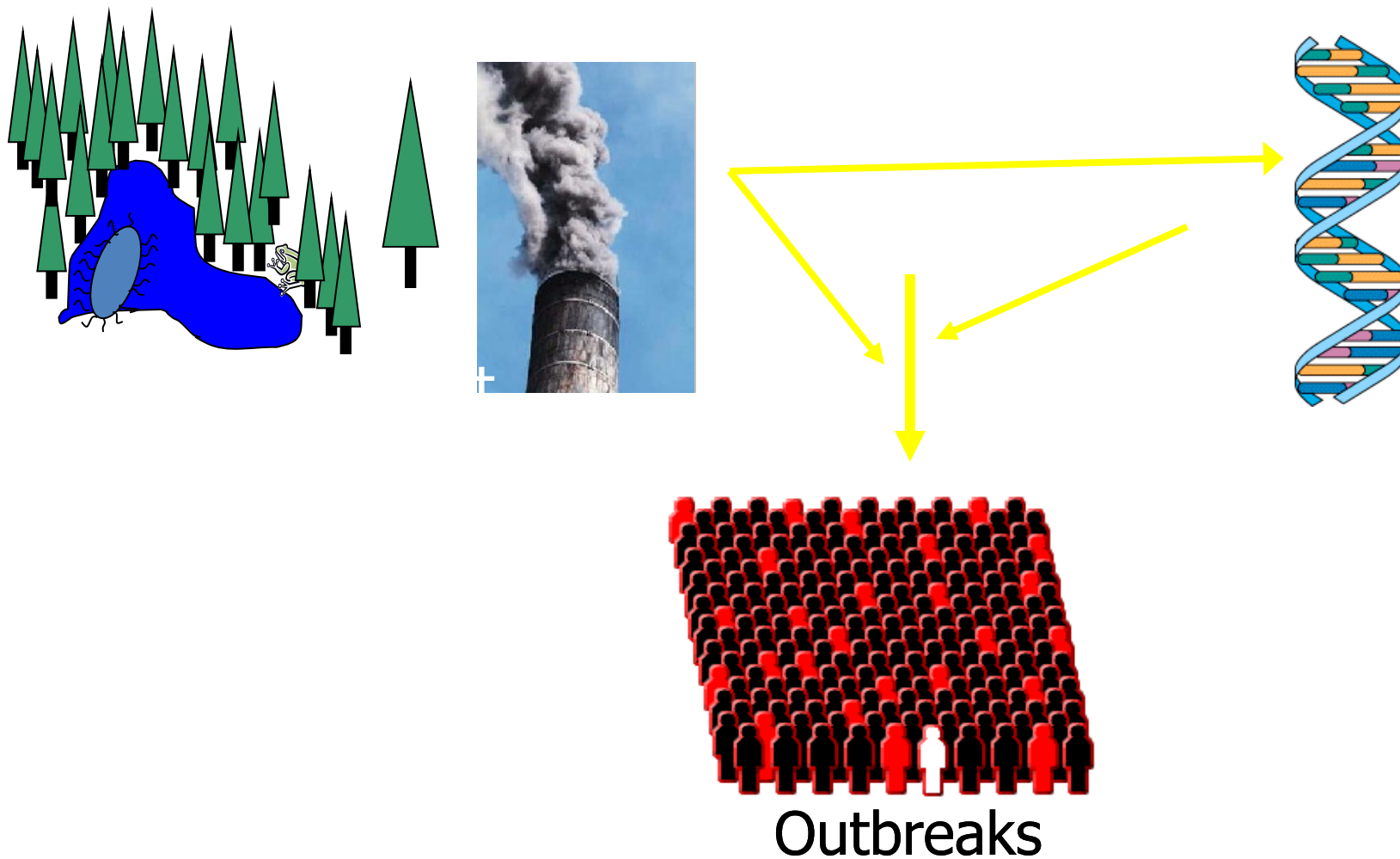


“MOLECULAR EPIDEMIOLOGY”

1. Markery expozice (např. přítomnost onkogenního viru)
1. Markery dávky (např. titr viru)
2. Markery interní dávky (např. množství DNA aduktů s karcinogenem)
1. Markery biologicky účinné dávky (např. somatické mutace p53)
1. Markery strukturního a funkčního poškození
2. (např. chromozomální aberace)
3. Markery vnímavosti (např. SNP genů metabolických enzymů)
1. Diagnostické markery
2. Prognostické a prediktivní markery

Schulte PA. *Molecular Epidemiology: Principles and Practices*. San Diego: Academic Press; 1993.

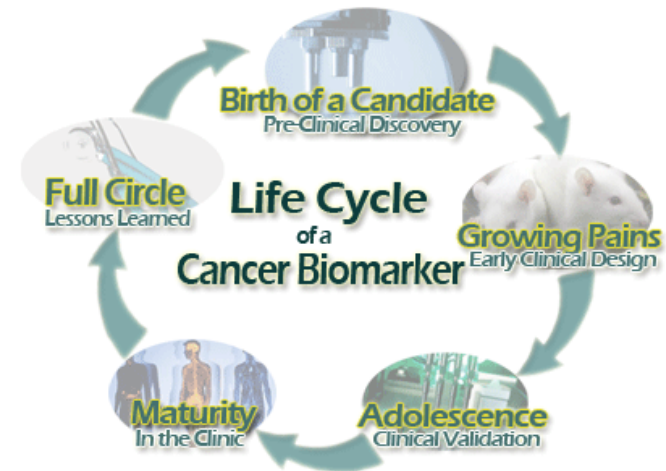
# Interakce molekulárních faktorů a prostředí



# Molekulární epidemiologie nádorových onemocnění

## Kategorie markerů používaných v epidemiologických studiích nádorových onemocnění:

1. Markery expozice (např. přítomnost onkogenního viru)
2. Markery dávky (např. titr viru)
3. Markery interní dávky (např. množství DNA aduktů s karcinogenem)
4. Markery biologicky účinné dávky (např. somatické mutace p53)
5. Markery strukturního a funkčního poškození (např. chromozomální aberace)
6. Markery vnímavosti (např. SNP genů metabolických enzymů)
7. Diagnostické markery
8. Prognostické a prediktivní markery



# Molekulární epidemiologie nádorových onemocnění

## Markery expozice – Příklad HPV a karcinomu děložního hrdla

100 podtypů HPV, a pouze limitovaný počet má onkogenní potenciál  
Vybrané podtypy HPV jsou asociovány s dlaždicobuněčným karcinomem, adenokarcinomem, dysplazií cervixu penisu, anu, vaginy, a vulvy.

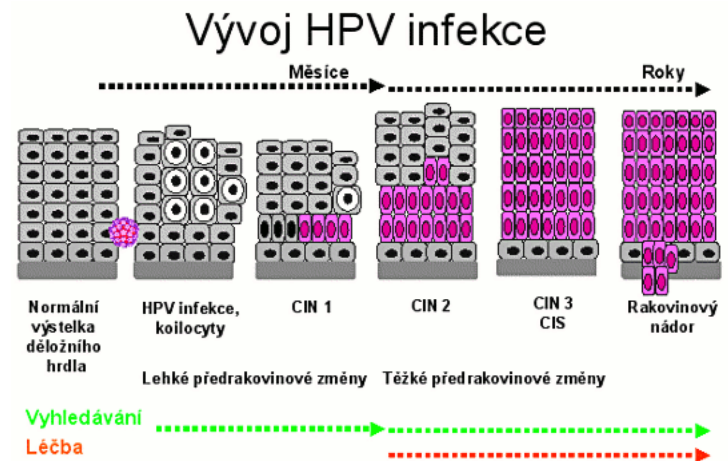
HPV je detekován v 95% až 100% vzorků cervikálního karcinomu -> a stal se tak „nutnou příčinou“ nádorů děložního hrdla.

Ženy s vyšším počtem kopií HPV viru v cervikální tkáni více rizikové (**marker dávky**).

Virapap (OR, **6.3**; 95% CI, 3.4-11.6)

Southernův přenos (OR, **16.3**; CI 7.7-34.4)

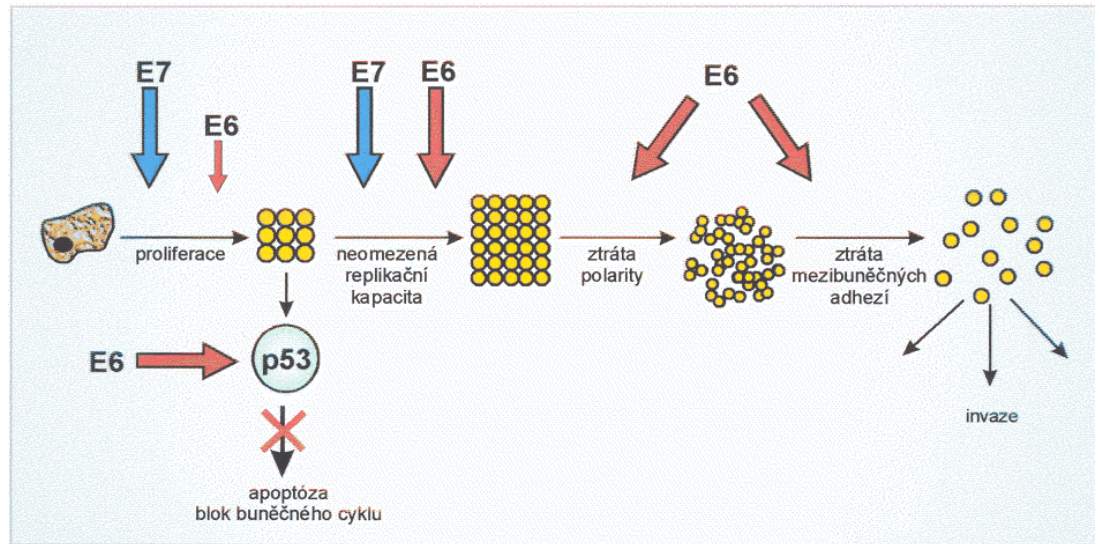
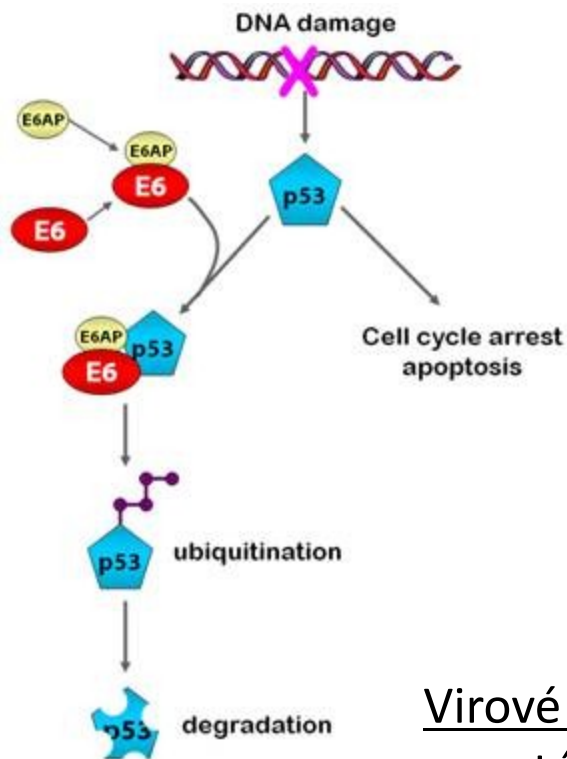
PCR (OR, **24.3**; CI, 14.4-41.0)



**Nutnost citlivých molekulárně biologických metod!!!!**

# Molekulární epidemiologie nádorových onemocnění

## Markery expozice – Příklad HPV a karcinomu děložního hrdla



Virové onkoproteiny E6, E7: stimulují proliferaci, inhibují apoptózu, zvyšují replikační potenciál, mění morfologii buněk, indukují maligní fenotyp

# Molekulární epidemiologie nádorových onemocnění

## Markery biologicky účinné dávky - somatické mutace

### *Ultrafialové záření indukuje mutace p53 a nádory kůže*

Klasická epidemiologie již dávno asociovala sluneční záření z nádory kůže, molekulární epidemiologie identifikovala specifické vlnové délky slunečního záření, které jsou karcinogenní (290-320 nm – UVB) a vedou k mutacím nádorového supresoru p53



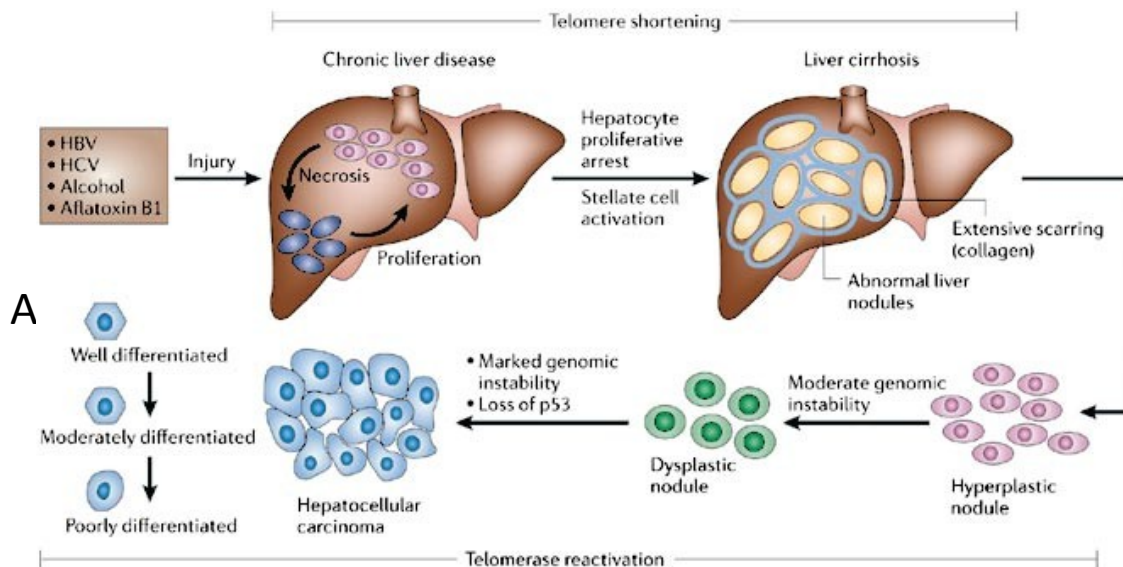
# Molekulární epidemiologie nádorových onemocnění

## Markery biologicky účinné dávky - somatické mutace

### Aflatoxin, hepatitidy a hepatocelulární karcinom

Nejrizikovější faktory pro hepatocelulární karcinom jsou expozice aflatoxinu, hepatitida B nebo C a alkohol. Aflatoxin je mykotoxin objevující se v kontaminovaných potravinách jako jsou ořechy, buráky, kukuřice a další. Aflatoxin tvoří adukty s DNA v moči a séru, které slouží jako biomarkery biologicky účinné dávky v epidemiologických studiích. Prospektivní studie

ukazují, že se hladiny těchto komplexů zvyšují během rozvoje hepatocelulárního karcinomu.



Expozice aflatoxinu vede ke klasické transverzi G:C na T:A ve třetí bázi kodonu 249 p53.



# Molekulární epidemiologie nádorových onemocnění

## Somatické mutace v nádorovém supresoru p53

Carcinogen Exposure	Neoplasm	Mutation
Aflatoxin B <sub>1</sub>	Hepatocellular carcinoma	Codon 249 (AGG → AGT)
Sunlight	Skin carcinoma	Dipyrimidine mutations (CC → TT) on nontranscribed DNA strand
Tobacco smoke	Lung carcinoma	G:C → T:A mutations on non-transcribed DNA strand (frequently codons: 157, 248, and 273)
Tobacco and alcohol	Carcinoma of the head and neck	Increased frequency p53 mutations (especially codons 157 and 248)
Radon	Lung carcinoma	Codon 249 (AGG → ATG)
Vinyl chloride	Hepatic angiosarcoma	A:T → T:A transversions

# Molekulární epidemiologie nádorových onemocnění

## Markery vnímavosti (Zárodečné mutace, hereditární syndromy)

**Penetrance** popisuje podíl jednotlivců nést zvláštní variaci genu to také vyjádřit spojovanou vlastnost.

Geny s vysokou penetrancí = rodinné studie, geny s nízkou penetrancí = asociační studie

### Germ-line mutace

#### Geny s vysokou penetrancí

**Karcinom prsu** (BRCA1, BRCA2, ATM-radiosensitivity, PTEN-Cowden syndrom, TP53-Li Fraumeni Syndrom)

#### Kolorektální karcinom

Lynchuv syndrom (MMR genes)  
Mikrosatelitní nestabilita (MSH2, MLH1,...)

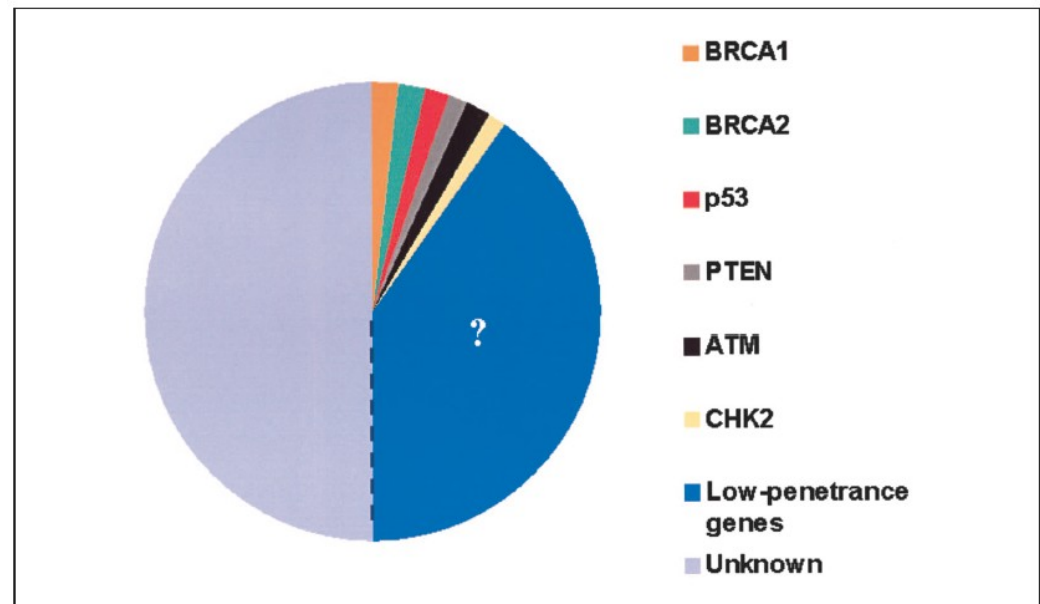


FIGURE 2 Proportion of Breast Cancer Attributable to Known and Unknown Germline Genetic Mutations.

# Molekulární epidemiologie nádorových onemocnění

## Markery vnímavosti (metabolismus karcinogenů, detoxifikační geny)

*Téměř všechny karcingeny vyžadují aktivaci metabolickými enzymy*

*Zároveň existují detoxifikační enzymy, které umožňují jejich odbourávání*

Při tepelné úpravě masa vznikají heterocyklické aminy a polycyklické aromatické uhlovodíky (PAH)

Např. heterocyklické aminy jsou metabolizovány

## N-acetyltransferázou 2 (NAT2)

**SNP v kódující sekvenci NAT2 vede k tvorbě proteinových variant dělicích bělošskou populaci**

**na pomalé acetylátory (55%, rychlé acetylátory 45%)**

*Pomalí acetylátoři konzumující velké množství červeného masa mají významně vyšší riziko CRC.*

## Take home

Epidemiologie, základní pojmy, kauzalita

Bradford Hillova kriteria kauzality

Faktory ovlivňující určení kauzality

Epidemiologické studie – dělení, přehled

Studie průřezové +příklad

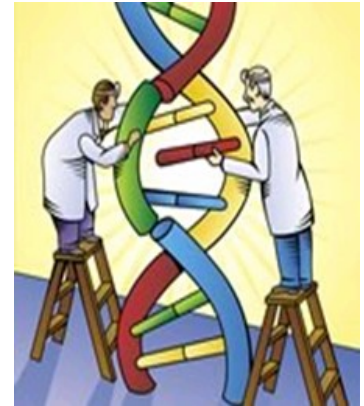
Studie případ-kontrola + příklad

Studie kohortové +příklad

Síla vztahu (asociace), statistické testování hypotéz

Definice molekulární epidemiologie (interakce molekulárních faktorů a prostředí)

Molekulární epidemiologie nádorových onemocnění (kategorie markerů)



## Náplň příští přednášky

Molekulární farmakologie I – cílená léčba – vývoj nových léčiv = identifikace nových molekulárních cílů, vysokovýkonný screening, tkáňové kultury, transgenní zvířecí modely, poměr rizik a prospěchu, ekonomická a etická hlediska při výběru identifikovaných cílů a vývoji nových léčiv

Molekulární farmakologie II – principy biologické léčby – monoklonální protilátky – příprava monoklonálních protilátek a rekombinantních proteinů, nízkomolekulární inhibitory – racionální design léčiv, siRNA, mikroRNA – tlumení genové exprese na post-transkripční úrovni, transport léčiv (lipozomy, imunoglobuliny, nanočástice a supramolekulární systémy)

## Dotazy?

