

Epidemiologické studie

Úvod do metodiky

Doc. MUDr. Jan Šimůnek, CSc.

Ústav preventivního lékařství

30. listopadu 2009

Základní pojmy

Epidemiologie

- Analyzuje vztahy mezi charakteristikami populačních skupin
- V případě zdravotnické e. analyzuje riziko a zdraví
- I studie typu: „Je výuková metoda A lepší/stejná/horší než výuková metoda B“ jsou epidemiologickými studiemi
- Analyzované veličiny chápe jako soubory *znaků*
- Pro analýzu využívá *matematickou statistiku*

Populační skupiny

Většinou nejsme schopni analyzovat skupinu jako celek, ale provádíme různé typy *výběrů*.

Výběry jsou extrémně náročné, silně ovlivňují výsledek studie, je třeba si na ně dát pozor i v případech *statistikace*

V zásadě musí být zajištěna rovnocennost pravděpodobnosti zařazení do výběru pro všechny členy populace, na něž se výsledek bude vztahovat.

Statistikace

Takto se nazývá (vy/zne)užití statistiky k podávání polopravdivých či zcela nepravdivých informací, argumentovaných „je statisticky prokázáno, že ...“
Nejčastějšími způsoby statistikace je zavádějící grafická presentace dat (je nutno si dát pozor na všechny grafy, které nemají souřadnice na nule), použití nevalidních statistických testů, ale lze se setkat i se statistikami na bázi zcela zfalšovaných vstupních dat.

Znaky

Znaky musejí být jednoznačně definované a musí být zajištěno, že jeden a tentýž jedinec dostane za všech okolností (třeba u různých hodnotitelů, v různých laboratořích apod.) tutéž hodnotu.

Typy znaků

číselné Nějaká číselná hodnota

pořadové Semikvantitativní hodnoty, umožňující seřazení, nemusejí být proporcionální

kvalitativní Hodnoty, které nelze rozumně dát do pořadí (např. krevní skupina). Zvláštní kategorie – znak **ANO/NE** má význam pro vznik tabulek 2×2 , hodnotitelných Fisherovým exaktním testem.

Typy studií

popisná v nejjednodušší formě popisuje výskyt určitých jevů v populaci. Tyto studie mají zásadní význam u sledování dosud neznámých jevů, kdy dle jejich výsledků plánujeme vyšší typy

analytické – viz příští slajd

intervenční předchozí studie popisují existující stav, intervenční do něj zasahují a hodnotí výsledky zásahu (analogie pokusů na zvířatech, které se hodnotí stejnými metodami); téměř veškeré testování nových léků a léčebných metod se děje intervenčními studiemi

metaanalýzy – viz o dva slajdy dále

Analytické studie

Analytické studie se snaží analyzovat vztahy mezi zjištěnými veličinami, případně jsou založeny tak, aby tato analýza byla dobře proveditelná. Hranice vůči k předchozímu typu studií je věcí domluvy: Budeme-li počítat z naměřených hodnot BMI a WHR indexy a sledovat jejich výskyt, tak půjde ještě patrně o studii popisnou, ale např. analýza který z těchto indexů více koreluje s prvky životního stylu nebo znaky rizika, třeba nad identickými daty, už je jednoznačně analytickou studií. Studie, které analyzují stav k určitému momentu jsou *průřezové*

Zvláštní typy analytických studií:

case-control studie případů a kontrol – u vzácných chorob, ke každému novému případu se zavádí jedna nebo více kontrol a analyzují se rozdíly a shody těchto dvou skupin

časově závislé analytické studie (následující slajd)

Časově závislé analytické studie

Retrospektivní

Využívají data z osobní anamnézy prováděné formalizovaným způsobem (porovnatelnost). Jsou zatíženy stejnými problémy jako osobní anamnéza obecně. Výhodou je jednorázový sběr dat, který je levnější, organizačně snazší a umožňuje zachovat anonymitu respondentů.

Prospektivní

Sledují určitou skupinu, kohortu (odtud alternativní název: *kohortové*) po určité období a průběžně sbírají data. Jsou extrémně finančně a organizačně náročné, mohou být zkresleny odpadnutím respondentů, odmítajících dlouhodobou spolupráci.

Metaanalýzy

Samy nemají své analyzované respondenty. Provádějí analýzu většího počtu studií na stejné téma, snaží se vysvětlit rozdíly nebo rozpory mezi nimi.

Módnost metaanalýz

Metaanalýzy potřebují kvalitní informace o studiích, které zahrnují. Určitá „módnost“ vede k tlaku na vytváření metaanalytických studií i nad pracemi, které jsou popsány nedostatečně a seriózně je analyzovat nelze. Na kvalitu vstupních dat je nutno dávat pozor!

Vytváření dotazníků, formulářů

kancelářské editory výhoda – zvládne každý, jsou všude,
nevýhoda – nelze v nich vytvořit typograficky
korektní dokument

T_EX, L_AT_EX výhoda – provozovatelné i na velmi slabých
počítačích, velmi kvalitní výsledek, nevýhoda –
musí se u toho myslet

profesionální DTP výhoda – jako předchozí, nevýhoda – jako
předchozí + cena

Zpracování a analýza dat

Počítačové programy

- kancelářské editory – jejich „pokračování“ Statistika pro windows
- Epi Info
- R
- databázové programy

Kancelářské editory

Jsou k dispozici „všude“.

Mají „zadrátovány“ základní statistické funkce.

Problémy

- Uzavřený formát – „životnost“ dat (výjimka Open Office org.). Zejména v akademickém prostředí je nutnost dlouhodobé dostupnosti dat
- Vypočtou cokoli nad prakticky jakýmkoli daty, i naprosté nesmysly, a bez jakéhokoli varování
- Mnohdy automaticky konvertují data do nesmyslné či nepoužitelné podoby

Epi Info

Program WHO původně pro operační systém DOS, nyní existuje i verze pro windows.

Verzi pro DOS lze spouštět v DOSBoxu, aplikaci emulující DOS na řadě OS.

Systém „pro blbý“, opatřený řadou varovných hlášení když data nebo výsledky nejsou kompatibilní s použitými testy.

Velmi jednoduchý formát dat, umožňující s nimi snadno pracovat dalšími programy.

R

Existuje profesionální statistický systém **S**, **R** je jeho free klonem.

Vyžaduje základní znalosti programování (spíše pochopení „o co se při tom jedná“, než znalost konkrétních programovacích jazyků).

Umožňuje i velmi složité statistické výpočty s grafickými výstupy a je dobře dokumentovaný.

Metoda volby v případě, kdy potřebujeme nějaký test, který není v Epi Info.

Databázové systémy

Umožňují především zpracování dat formátovaných jinak než do klasické tabulky (R to umí taky, ale „tabulka“ je jeho základní prostředek).

Poskytují rozsáhlé možnosti programování, případně programovat nad databází v dalších programovacích jazycích než je jazyk samotné databáze.

Interpretace výsledků 1

Epidemiologické studie většinou poskytnou informace o „statistické významnosti“ mezi některými znaky.

Statistická významnost

„Statistická významnost“ představuje vyhodnocení pravděpodobnosti, že daná čísla vznikla ryze náhodně. Nepředstavuje *záruku*, že vztah náhodný není. Na to je třeba dávat pozor u „heuristických“ studií typu „nahromadíme hodně dat a poté vše analyzujeme“. Máme-li dostatečně rozsáhlou databázi (počet znaků), pak po jejím naplnění ryze náhodnými čísly je vysoká pravděpodobnost, že vztahy mezi některými znaky budou „statisticky významné“. Do jisté míry tuto skutečnost reflektují jednostranné a dvoustranné varianty některých statistických testů.

Interpretace výsledků 2

Náhodná koincidence

Na tuto variantu musíme vždy pomýšlet při nějakých „záhadných“ výsledcích.

Zprostředkovaný vztah

Vztah mezi veličinami A a B může být zprostředkován jejich vztahem k veličině C, kterou jsme nezjišťovali (nebo ani zjistit neumíme), přitom se vzájemně neovlivňují.

Co závisí na čem

Výsledkem statistiky je vztah mezi znaky, není možno rozlišit příčinu od následku. *Tuberkulózní krávy dávají statisticky více mléka. Realita: Protože se kráva nadprůměrnou doživostí více vyčerpá, snadněji se nakazí TBC.*

Presentace dat 1

Grafy a schémata statistický software

kancelářské editory, případně Statistika pro windows tváří se jednoduše, jakmile chceme něco mimo předpřipravená schémata, obrovské problémy

EpilInfo pro windows o něco lepší možnosti, ale opět jsme vázáni na předpřipravené vzory

- R Má rozsáhlé možnosti, není s ním zcela triviální práce, dají se tam domalovávat i primitiva (čáry, body, mnohoúhelníky, ovály)

Presentace dat 2

Grafy a schémata grafický software

inkscape vektorový editor, nativní formát je SVG, lze do něho zasahovat i textovým editorem

gimp bitmapový editor, má makrojazyk, prostřednictvím kterého by se daly dělat i jednoduché grafy. Masochistům vřele doporučuji!

metapost vektorová grafika nad velice silným makrojazykem (odvozeným od $\text{T}_{\text{E}}\text{X}$), patrně nejlepší 2D grafika vůbec

POVRay 3D grafika, je možné buď přímo nadefinovat objekty v jazyce POVRaye, nebo použít editor Blender a přes POVRay jen renderovat scény. Patrně nejlepší 3D grafika, která je zdarma.

Děkuji za pozornost