

PACIENTI SE SEXUÁLNĚ PŘENOSNÝMI NEMOCEMI S RIZIKOVÝM CHOVÁNÍM V PRAŽSKÉ POPULACI 2007-8 (PRŮŘEZOVÁ STUDIE)

Jaromír Běláček, Ivana Kuklová, Petr Velčevský,
Ondřej Pecha, Marek Novák

Anotace

V letech 2008 - 2010 byla na oddělení BioStat ÚBI 1.LF UK a VFN v Praze provedena řada statistických analýz u osob vyšetřovaných s podezřením na získané sexuálně přenosné onemocnění (Sexually Transmitted Diseases - STD), které byly anonymně dotazovány prostřednictvím sady standardizovaných anketních otázek na základní demografické charakteristiky a behaviorální (sexuální) zvyklosti.

U většiny osob (přesněji 301/61 mužů/žen) ze souboru N=488 dotazovaných v letech 2007-8 byla diagnostikována některá STD z následujícího seznamu: syfilis primaria (21/0); syfilis secundaria recens (54/23), syfilis secundaria latens (12/8); gonorea(122/23); chlamydiální infekce (78/9); urogenitální infekce vyvolaná *Mycoplasma hominis* (2/4); *Ureaplasma urealyticum* (31/12); hepatitis C (4/1); HIV (7/0); infekce vyvolaná papilomaviry (24/0); u 28 pacientů se syfilis byla diagnostikována ještě jiná STD. Výběrový soubor je typický vysokým podílem osob s rizikovým sexuálním chováním (22% homosexuálů - mužů; více než 9% vyšetřovaných „nepoužívá nikdy kondom“; 21% uvedlo „první pohlavní styk ve věku do 15 let“; 10% dotazovaných přiznalo „více než 10 partnerů za posledních 12 měsíců“; 5% „poskytlo sexuální službu za úplatu“ atd).

V rámci této studie jsme se pokusili formálními statistickými prostředky specifikovat, které z výše naznačených souvislostí mezi sledovanými proměnnými a STD nemocemi se indikují jako statisticky významné. Protože některé proměnné mohou být s nemocemi v kauzálním vztahu a některé s nimi pouze koexistovat, použili jsme pro analýzu několika zdánlivě různých statistických metod (matice párových korelačních koeficientů, testy chí-kvadrát v kontingenčních tabulkách, ROC analýzy). Prezentované výsledky byly dosaženy a vznikly za podpory GP NS10292 (MZ0/NS, 2009-2011).

Klíčová slova:

Faktorová a korelační analýza, sexuálně přenosné nemoci (STD), rizikové sexuální chování, rizikové skupiny, ROC analýza

1. Úvod

Zvýšení incidence a prevalence nebezpečných sexuálně přenosných nemocí, jako je syfilis, kapavka, chlamydiální infekce, HIV/AIDS a další STD (Sexually Transmitted Diseases), zaznamenává v posledních letech u pražské populace nebývalý boom (bohužel i z řady objektivních příčin). Dotazování pacientů pražského STD centra infikovaných některou ze získaných STD na jejich behaviorální a sexuální zvyklosti probíhá na Dermatovenerologické klinice 1.LF UK Praha periodicky již řadu let (viz např. v [1]). Aktualizovaný dotazník

obsahuje t.č. cca 25 strukturovaných otázek, které zahrnují informace demografické (pohlaví, věk, národnost, rodinný stav, vzdělání, zaměstnání respondentů), epidemiologické (jaké zdravotní obtíže přivádějí pacienta; kdo jej poslal na vyšetření apod.), dotazy na sexuální chování (sexuální orientaci, praktiky, počet sexuálních partnerů za posledních 12 měsíců) až po frekvenci a formy používání (či nepoužívání) antikoncepce včetně doprovodných rizikových faktorů (drogy, kouření, konzumace alkoholu atd.).

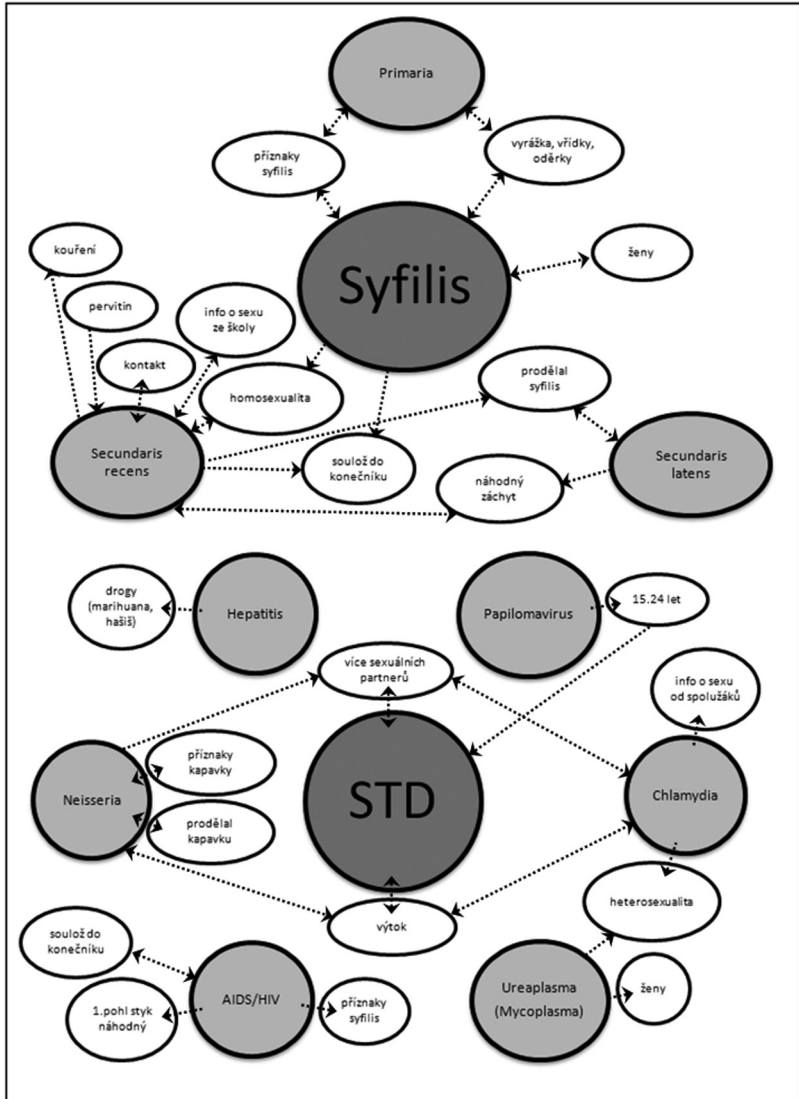
Cílem této studie bylo ověřit, zda event. jakými formálně statistickými prostředky lze výše jmenované ukazatele považovat za markery laboratorně diagnostikovaných STD nebo dokonce za latentní (skryté) prediktory pro výskyt těchto nemocí v populaci.

2. Statistické metody

Za účelem systematické analýzy prezentované v rámci tohoto příspěvku byly všechny analyzované proměnné nejprve překlasifikovány na dvouhodnotové (dichotomické) veličiny: Řada otázek v dotazníku, jmenovitě proměnné klasifikující lékařem verifikované STD diagnózy, byly dichotomické již ze své vlastní podstaty (ať byly dotazovány kupř. ve formě '1-Ano'; '2-Ne'; '3-Nevím' anebo "s možností více odpovědí" tzv. multiresponse). Vícekategoriální odpovědi z dotazníku byly převedeny na nula-jedničkové buď prostřednictvím tzv. dummy proměnných ('1~kategorie zodpovězená jako Ano'; '0~ve všech ostatních případech') anebo byly u vybraných „ordinálních“ proměnných, jako je např. „věk“ dotazovaných, rozděleny na dvě hodnoty ('1-Low'; '2-High') vhodně zvolenou agregací.

Pro takto vytvořenou sestavu cca 90 dvouhodnotových proměnných byly vypočteny matice všech párových korelačních koeficientů – za celý soubor respondentů (N=488), za všechny pacienty se syfilis (N=87/31), za všechny pacienty s nálezem jiné STD (N=235/37) a za zbytkovou skupinu kontrol (N=107/21). Následně byla provedena formální analýza statistické významnosti korelačních koeficientů; u korelačních koeficientů pro všechny dvojice různých proměnných proti nulové hypotéze ($H_0: r=0$), u vybraných skupin nemocných vůči kontrolám proti hypotéze o rovnosti korespondujících korelačních koeficientů ($H_0: r_{STD} = r_{CONTROLS}$).

Vstupní sestava všech dichotomických proměnných (o rozsahu N=488) byla podrobena faktorové analýze (viz [3], str. 663-680) s cílem: Co nejobektivněji identifikovat počet a hlavní skupiny tzv. skrytých/latentních resp. dobře interpretovatelných faktorů, které by měly schopnost vysvětlit vnitřní strukturu korelační matice resp. shledaných signifikancí „s dostatečně velkou přesností“ (měřeno procentem vysvětlené variability této matice). Speciální pozornost při této analýze se zaměřila na faktory obsahující jednotlivé STD. V tomto smyslu byla provedena analýza citlivosti (senzitivity a specifity) a predikční schopnosti sexuálně-behaviorálních ukazatelů vůči indikovaným STD nemocem prostřednictvím standardních ROC analýz (viz [3], str. 1717-1720). Naše hlavní zjištění jsou ve zjednodušené podobě prezentována v následující stati a prostřednictvím jednoduchého grafického schématu.



Obrázek 1 - Grafické schéma statisticky významných kauzálních závislostí mezi diagnostikovanými STD a behaviorálními charakteristikami

2. Výsledky

2.1 Struktura faktorových zátěží výběrového souboru

Korelační analýza provedená na úrovni celého souboru (N=488) respondentů prokázala vysokou statisticky významnou závislost mezi mnoha z 95 vstupních dichotomických proměnných. Kdybychom si přáli vysvětlit variabilitu této korelační matice ze 75%, museli bychom použít 35 nejvýznamnějších faktorů; pro aproximaci na úrovni 50% by jich bylo zapotřebí 17. Od této hranice jsme postupně snižovali počet faktorů o jedničku, až jsme se konečně zastavili na počtu 9 interpretovatelných, které vysvětlují cca 35% variability. Přestože výsledky nejsou zcela jednoznačné, lze (vzájemně nezávislé latentní) faktory přibližně identifikovat následujícími sestavami proměnných:

- I: PŘÍZNAKY SYFILIS (pacienta přivedly k vyšetření nebo v minulosti prodělal syfilis); diagnostikovaná syfilis (primaria nebo secundaria-recens) nebo HIV; vřídky, oděrky na genitálu;
- II: DROGY (uváděn pervitin, heroin, marihuana, hašiš; injekční aplikace); diagnostikovaná hepatitis; přiznaná sexuální služba za úplaty;
- III: HOMOSEXUÁLNÍ PRAKTIKY (pohlavní styk do konečníku, do úst, nekointimální praktiky); přednost stykům se stejným pohlavím; více než 4 partneři za posledních 12 měsíců; v minulosti prodělal kapavku;
- IV: DIAGNOSTIKOVANÁ SYFILIS LATENS (náhodný záchyt nebo na vyšetření posláni gynekologem); obvykle žena; gravidní nebo cizí národnosti;
- V: VĚK 15-24 LET (svobodný/á; student/ka); faktor asociován s diagnostikovaným výskytem humánního papilomaviru;
- VI: PŘÍZNAKY KAPAVKY (pacienta přivedly k vyšetření); diagnostikovaná *Neisseria gonorrhoe*;
- VII: PROFESNĚ MĚNĚ KVALIFIKOVANÉ OSOBY (dělníci, prodavači apod.); obvykle nepoužívají žádnou antikoncepci (kondom nikdy nebo zřídka); první pohlavní styk ve věku 16-17 let; kuřáci;
- VIII: VYSOKOŠKOLÁCI (odesláni na vyšetření praktickým lékařem); první pohlavní styk v rámci partnerského vztahu; příležitostně pít alkoholu;
- IX: PŘÍZNAKY JINÉ STD (diagnostikovány příznaky *Chlamydia trachomatis*, *Mycoplasma hominis* nebo *Ureaplasma urealyticum*; na vyšetření odesláni urologem); středoškoláci s maturitou; častější výskyt u zaměstnanců z oblasti erotických a zábavních služeb.

Ne všechny výše uvedené, jmenovitě např. skupiny V, VII nebo VIII, jsou bezprostředně vázány na konkrétně diagnostikovanou STD nebo je lze přímo interpretovat jako latentní/skrytý zdroj pro výskyt některé STD (výběrový soubor obsahuje m.j. i zbytkovou skupinu N=128 „kontrol“).

2.2 Výsledky kauzálních (ROC) analýz

Hlavní výsledky ROC analýz jsou vyjádřeny na dvou níže uvedených schématech. Šipky vyjadřují statisticky významné kauzální relace pro $p \leq 0,05$ (Ve směru „od“ na grafu vyznačených D_g jde vlastně o analýzu senzitivity

a specifický prostřednictvím statistické významnosti ukazatelů AUC vůči referenční hodnotě 0,5 – viz [3], str.1720; ve směru „k“ vyznačeným STD jde naopak o formálně ověřenou schopnost vybraných behaviorálních či dalších anketně dotazovaných indikátorů předpovědět výskyt některé na grafu uvedené STD prostřednictvím tzv. pozitivních a negativních prediktivních hodnot – viz také v [4], str. 138.)

3. Diskuse

Výše zvolený metodický přístup umožňuje jednak univerzálnější aplikaci metod mnohorozměrné statistické analýzy i následnou snazší orientaci ve statistické významnosti výsledků mnohorozměrných výstupních statistických sestav. Současně je bližší i snahám medicínských specialistů o maximální srozumitelnost výsledků, protože přirozeně pracuje s nejjednoduššími matematicko-statistickými objekty (procenta, čtyřpolní tabulky, elementární podmíněné pravděpodobnosti). Neumožňuje však přímo proniknout (věcně i formálně) do vnitřní struktury analyzovaných dat, což je nutné spojeno s řadou speciálních kombinačních třídění, např. podle demografie respondentů (pacientů i náhodně zachycených kontrol), zjištěných STD diagnóz a specifických behaviorálních/sexuálních zvyklostí. Některé z analýz tohoto typu byly již provedeny (např. v [1]), některé další (detailnější) budou ještě následovat.

4. Poděkování

Tato práce byla vytvořena za podpory *Projektu NS10292 – Vývojové trendy a prevalence syphilis, gonorey, chlamydiální infekce a HIV/AIDS u osob s rizikovým sexuálním chováním; behaviorální, klinické a epidemiologické charakteristiky pacientů STD centra (2009-2011, MZO/NS)*; příjemce: 1.LF UK Praha, řešitel: MUDr. Ivana Kuklová, CSc., další řešitel: RNDr. Jaromír Běláček, CSc.;spolupříjemce: VFN Praha, řešitel: MUDr. Martina Kojanová, M.D., další řešitelé: RNDr. Marcela Draždáková, MUDr. Roman Trýzna, M.D. – viz [2].

Literatura

- [1.] Kuklová I-Velčevský P-Kojanová M-Kašátková V-Trýzna R-Pánková R-Sedláková K-Běláček J: *Analýza příčin stoupající incidence syfilidy v pražské populaci. Čes-slov Derm, 2009, 84/6, str. 350-355*
- [2.] *Projekt NS10292 [online]. Informační systém výzkumu, experimentálního vývoje a inovací. [cit. 2010-02-08]. Dostupný z WWW : <<http://www.isvav.cz/projectDetail.do?rowId=NS10292>>*
- [3.] *SPSS (2007): 17.0 Command Syntax Reference, Chicago, IL 60606-6412, str. 663-680, 1717-1720, viz www.spss.cz*
- [4.] *Špunda M.-Dušek J.(eds.): Zdravotnická informatika. Kap. 9.1. Karolinum, skripta 1.LF UK Praha, 2007, ISBN 978-80-246-1378-9*

Kontakt:

RNDr. Jaromír Běláček, CSc,

Mgr. Ondřej Pecha, Ph.D.

Ing. Marek Novák

Ústav biofyziky a informatiky

1.LF UK a VFN Praha

Salmovská 1, 121 08 Praha 2

tel: +420 22496 5706 nebo 1.5797

e-mail: jaromir.belacek@lf1.cuni.cz

<http://biof.lf1.cuni.cz/biostat.html>

MUDr. Ivana Kuklová, CSc.

MUDr. Petr Velčevský

Dermatovenerologická klinika

1.LF UK a VFN Praha

U nemocnice 2, 128 08 Praha 2

tel: +420 22496 6731 nebo 1.2406

e-mail: ivana.kuklova@lf1.cuni.cz

www.lf1.cuni.cz