

ÚVOD DO PROBLEMATIKY ANALÝZY PŘEŽÍVÁNÍ INTRODUCTION INTO THE SURVIVAL ANALYSIS

KOUKALOVÁ H. (1), PECEN L. (2)

MASARYKŮV ONKOLOGICKÝ ÚSTAV, BRNO (1)

ÚSTAV INFORMATIKY A VÝPOČETNÍ TECHNIKY AKADEMIE VĚD ČESKÉ REPUBLIKY, PRAHA (2)

Úvod

Hlavním cílem boje proti zhoubným nádorovým onemocněním je snížení mortality. Toho může být dosaženo jednak snížením incidence, jednak zlepšováním léčebných výsledků. V tomto směru jsou cenným údajem informace o přežití, získané na základě dat onkologického registru (pokud možno populačního, ne pouze nemocničního) zpracované metodami analýzy přežívání.

Pojem přežití jako předmět statistické analýzy přežívání neznamená pouze dobu od vzniku onemocnění do úmrtí. Obecně se jím označuje délka časového intervalu od určitého okamžiku (např. počátek expozice rizikovým faktorům, datum stanovení diagnózy, ukončení terapie apod.) do výskytu očekávané události (např. vznik onemocnění, relaps, úmrtí). V nádorové epidemiologii bývá zpravidla sledována doba od stanovení diagnózy do úmrtí.

Popis metody

Metody analýzy přežití umožňují výpočet (odhad) **funkce přežití S(t)** definované vztahem

$$S(t) = P(T \geq t),$$

kde t je čas (ve stanovených jednotkách - měsících, letech,...)

T je délka přežití pacienta

P je označení pro pravděpodobnost

Funkce přežití je tedy pravděpodobnost, že délka přežití pacienta je větší nebo rovna času t, tj. že pacient přežije alespoň t let od stanovení diagnózy. Za t je možno dosadit libovolné, z klinického hlediska smysluplné číslo.

Při výpočtu (odhadu) funkce přežití času t, může u každého pacienta ve zpracovávaném statistickém souboru nastat právě jedna z následujících možností:

1. Pacient přežil dobu delší než t
2. Pacient zemřel do času t na sledovanou nemoc
3. Pacient zemřel do času t na jinou než sledovanou nemoc
4. Není známo, zda pacient v čase t žije nebo již zemřel, protože doba jeho sledování byla kratší než t. Např. proto, že se odstěhoval do zahraničí a od tohoto okamžiku jsou údaje o něm nedostupné, nebo, a to nejčastěji proto, že diagnóza byla stanovena před dobou kratší než t. Takto nekompletní **data** se nazývají **cenzorovaná**.

V případě, že u všech sledovaných pacientů nastaly pouze možnosti 1) - 3), tj. k dispozici jsou úplné údaje, jde o trivi-

ální záležitost. Funkci přežití $S(t)$ zde lze spočítat jako podíl počtu osob, které přežily čas t , k celkovému počtu osob v souboru. Tedy jestliže ze 100 pacientů se jich 20 dožije 5 let od stanovení diagnózy, pak $S(5) = 0,2$, neboli 20%.

V epidemiologii nebo obecně ve studiích, jež jsou z hlediska času i velikosti souboru rozsáhlejší, však jsou často využívány i údaje neúplné, cenzorované. Tyto není vhodné jednoduše ze zpracování vyřadit (možnost zavlečení zkreslení), ale současně kvůli nim není možno přesně spočítat funkci přežití $S(t)$. Tu lze pouze odhadnout, a to právě metodami analýzy přežití.

Jednou z nejpoužívanějších metod analýzy přežití je **metoda Kaplan-Meierova**, jež je maximálně věrohodným bodovým odhadem funkce přežití. Výpočet odhadu funkce $S(t)$ je zde dán vztahem

$$P(T \geq t_i) = P(T \geq t_0) * P(T \geq t_1 | T \geq t_0) * \dots * P(T \geq t_i | T \geq t_{i-1})$$

kde t_i je čas, obvykle $i = 0, 1, \dots$

T je délka přežití

i je označení pro podmínku

Odhad pravděpodobnosti přežití počátečního času t_0 je počítán z údajů o všech pacientech. Do výpočtu podmíněné pravděpodobnosti přežití alespoň doby t_i za podmínky, že doba života je alespoň t_{i-1} (zapsáno jako $P(T \geq t_i | T \geq t_{i-1})$), vstupují pouze údaje o pacientech, kteří v čase t_{i-1} žijí a jejichž doba sledování je alespoň t_{i-1} . Nezahrnují se údaje o pacientech, kteří do času t_{i-1} zemřeli, nebo o pacientech, jejichž doba sledování je menší než t_{i-1} . Tímto způsobem jsou využita data všech pacientů, a to tak, že ovlivňují výpočet právě tak dlouho, jak dlouho je pacient ve sledování.

Příklad: Vztah pro odhad pravděpodobnosti přežití pacienta alespoň 4 roky od stanovení diagnózy:

$$P(T \geq 4 \text{ r.}) = P(T \geq 1 \text{ r.}) * P(T \geq 2 \text{ r.} | T \geq 1 \text{ r.}) * \\ P(T \geq 3 \text{ r.} | T \geq 2 \text{ r.}) * P(T \geq 4 \text{ r.} | T \geq 3 \text{ r.})$$

Závěrem zbývá dodat, že **kvalifikované analýze přežívání musí vždy předcházet důkladná příprava dat spočívající v jejich kontrole a vyloučení záznamů, které by mohly tento ukazatel zkreslovat. Získané výsledky by měly být vždy interpretovány ve spolupráci statistika, onkologa a epidemiologa.**

Pozn.: Funkce přežití bývá nestatisticky mnohdy interpretována jako procento osob přeživších daný počet let od stanovení diagnózy. Přestože tato interpretace je správná pouze pro data bez cenzorů, je pro větší srozumitelnost tolerována a běžně se používá i v případě dat cenzorovaných.

Používané ukazatele

V souvislosti s přežíváním se lze nejčastěji setkat s následujícími typy tohoto ukazatele:

a) **absolutní přežití** (absolute, observed, overall survival) – očekávanou událost se rozumí úmrtí na jakékoliv onemocnění.

b) **přežití specifické dle diagnózy** (cause-specific survival) – očekávanou událost je pouze úmrtí na sledované onemocnění, přičemž úmrtí na jinou příčinu je považováno za cenzor (informace o úmrtí má v tomto případě stejnou hodnotu jako např. odstěhování pacienta do zahraničí). Hodnoty tohoto ukazatele jsou však závislé na kvalitě záznamů o příčinách úmrtí, jež je často sporná, resp. vykazuje vysokou variabilitu.

c) **relativní přežití** (relative survival) – poměr mezi absolutním přežitím u pacientů s danou nemocí a absolutním přežíváním v celé populaci. Přesněji: jde o vážený průměr těchto poměrů v podskupinách rozdělených podle pohlaví a věkových kategorií, s vahami obvykle proporcionálními struktuře dle věku a pohlaví populace postižené daným nádorem. Tento ukazatel, jehož hlavní výhodou je nezávislost na záznamech o příčinách úmrtí sledovaných pacientů, informuje o tom, „na kolik procent snižuje dané nádorové onemocnění pravděpodobnost přežití času alespoň t oproti celé populaci“. Tedy např. pokud bude hodnota 5letého relativního přežití mužů s nádorem plic ve stádiu I rovna 20, znamená to, že přežití takto postižených osob je oproti přežití celé populace sníženo na pouhých 20%. Pravděpodobně nejužívanější metodou výpočtu relativního přežití je metoda Hakulinenova (3), jež i přes určité interpretační problémy s ní spojené je asi tím nejlepším, co je v současné době na dostupná data možno aplikovat.

Závěr

Přežívání onkologických pacientů je jedním z hlavních kritérií hodnocení účinnosti screeningových projektů, diagnosticko-léčebných postupů i celých onkologických programů. Využívání tohoto ukazatele proto nabývá stále většího významu. Článek přináší základní informace o metodě a ukazatelích přežívání. Upozorňuje na nezbytnost pečlivé přípravy dat i kvalifikované interpretace výsledků. Podrobnější informace o problematice lze získat v uvedené literatuře.

Literatura

1. Berrino, F., et al. Survival of cancer patients in Europe : The EURO CARE study. Lyon : IARC, 1995. 463 p.
2. Estève, J., Benhamou, E., Raymond, L. Descriptive epidemiology. Lyon : IARC, 1994. 302 p.
3. Hakulinen, T. Cancer survival corrected for heterogeneity in patient withdrawal. Biometrics, 1982, vol. 38, p. 933-942.