

**Univerzita Palackého v Olomouci**

**Lékařská fakulta**

**ANALÝZA PŘEŽÍVÁNÍ  
– VYTĚŽOVÁNÍ INFORMACÍ Z CENZOROVANÝCH DAT**

Dizertační práce

**Mgr. Kateřina Langová**

Školitel: Prof. RNDr. Stanislav Komenda, DrSc.

Školící pracoviště: Ústav lékařské biofyziky, LF UP

**OLOMOUC 2008**

Prohlašuji, že jsem dizertační práci vypracovala sama pod odborným vedením  
Prof. RNDr. Stanislava Komendy, DrSc. a že jsem uvedla veškeré použité zdroje a literaturu.

Olomouc, 2008

.....

Děkuji svému školiteli Prof. RNDr. Stanislavu Komendovi, DrSc. za odborné vedení dizertační práce a odbornou pomoc v průběhu studia. Dále pak děkuji Prof. MUDr. Vítězslavu Kolkovi, CSc., Prof. MUDr. Vlastimilu Ščudlovi, CSc., MUDr. Luděku Raidovi, Ph.D., Doc. MUDr. Martinu Kócherovi, Ph.D. a Doc. MUDr. Jiřímu Gallo Ph.D. za laskavé poskytnutí klinických dat, která byla použita v této práci. Dík patří také Ing. Jiřímu Macháčkovi za pomoc při vytváření softwarové podpory.

## OBSAH

Seznam zkratek a symbolů .....	6
<b>1 ÚVOD .....</b>	<b>8</b>
1.1 Analýza přežívání.....	10
1.2 Cenzorování .....	10
1.3 Funkce přežití .....	13
1.4 Funkce hazardu .....	14
<b>2 ODHADY DISTRIBUCE PŘEŽÍVÁNÍ .....</b>	<b>19</b>
2.1 Kaplan-Meierův odhad distribuce přežívání.....	19
<b>3 POROVNÁNÍ DVOU NEBO VÍCE KŘIVEK PŘEŽITÍ.....</b>	<b>23</b>
3.1 Log-rank test pro dvě skupiny .....	23
3.2 Log-rank test pro více skupin .....	28
3.3 Wilcoxonův test pro porovnání dvou křivek přežití .....	31
3.3.1 $W$ statistika.....	31
3.3.2 Průměr a rozptyl statistiky $W$ .....	32
3.3.3 Výpočet statistiky $W$ a $\text{var}(W P, H_0)$ ve velkých vzorcích.....	34
3.3.4 Příklad.....	35
<b>4 COXŮV REGRESNÍ MODEL.....</b>	<b>39</b>
4.1 Coxův model proporcionálního hazardu .....	39
4.2 ML (maximum likelihood) odhady regresních koeficientů Coxova modelu .....	41
4.2.1 Likelihood funkce.....	41
4.2.2 ML odhady .....	42

4.3	Výběr Coxova modelu.....	43
4.4	Výpočet hazard ratio ( <i>HR</i> ) .....	45
4.5	Coxův model funkce přežití.....	46
5	MEDIÁN A PRŮMĚR V ANALÝZE PŘEŽÍVÁNÍ .....	48
5.1	Medián.....	48
5.2	Průměr.....	48
6	POČÍTAČOVÁ PODPORA .....	50
6.1	Standardní aplikace .....	50
6.2	Vlastní programové vybavení .....	50
7	PŘÍKLADY .....	57
7.1	Studie přežívání hematoonkologických pacientů .....	57
7.2	Studie přežívání pacientů s mnohočetným myelomem.....	64
7.3	Studie sledování průchodnosti stehenní tepny po subintimální angioplastice.....	74
7.4	Studie sledování přežívání endoprotéz u pacientů po operaci kyčelního kloubu...	83
8	ZÁVĚR.....	91
	Souhrn.....	92
	Summary .....	93
	Literatura .....	94
	Publikační činnost autora .....	97

## Seznam zkratek a symbolů

$e$  – Eulerovo číslo, přibližně 2,71

$e^x$  – exponenciální funkce se základem  $e$

$E(X)$  – střední hodnota náhodné veličiny  $X$

$F(x)$  – distribuční funkce

$f(x)$  – hustota rozdělení

$h(t)$  – funkce hazardu, riziková funkce

$h_0(t)$  – funkce základního hazardu

$HR$  – hazard ratio, poměr hazardů

$\hat{HR}$  – odhad hazard ratio

$\ln(x)$  – funkce přirozený logaritmus

$L(x)$  – věrohodnostní funkce

$LR$  – likelihood ratio, věrohodnostní poměr

$ML$  – maximum likelihood, maximální věrohodnost

$N$  – počet, četnost

$P(A)$  – pravděpodobnost výskytu jevu  $A$

$P(A | B)$  - podmíněná pravděpodobnost jevu  $A$  vzhledem k jevu  $B$

$PI$  – prognostický index

$S(t)$  – funkce přežití

$S_0(t)$  – základní funkce přežití

$\hat{S}(t)$  – odhad funkce přežití

$T$  - doba od zahájení pozorování jedince do výskytu určité události, doba přežití

$t$  - konkrétní hodnota pro  $T$

$\Delta t$  – přírůstek času

$\text{Var}(X)$  – rozptyl náhodné veličiny  $X$

$X_1, \dots, X_p$  -  $p$  nezávisle proměnných v Coxově modelu

$\alpha$  – parametr Weibullova rozdělení

$\beta_1, \dots, \beta_n$  - regresní koeficienty v Coxově modelu

$\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_n$  - odhady regresních koeficientů v Coxově modelu

$\Gamma(x)$  – funkce gama

$\delta$  – parametr logaritmicko-normálního rozdělení

$\varepsilon$  - náhodná veličina, která označuje, zda došlo k dané události ( $\varepsilon = 1$ ) či nikoliv ( $\varepsilon = 0$ )

$\theta$  - parametr věrohodnostní funkce

$\lambda$  - parametr exponenciálního a Weibullova rozdělení

$\mu$  - parametr (střední hodnota) normálního rozdělení

$\pi$  - Ludolfovo číslo, přibližně 3,14

$\Pi$  - součin

$\sigma$  - parametr (směrodatná odchylka) normálního rozdělení

$\Phi\left(\frac{x-\mu}{\sigma}\right)$  - distribuční funkce normálního rozdělení se střední hodnotou  $\mu$  a rozptylem  $\sigma^2$

# 1 Úvod

Problém analýzy času do výskytu nějaké události se objevuje v různých oblastech lidského bádání, například v medicíně, biologii, epidemiologii, inženýrství, ekonomii, sociologii a dalších. Ačkoliv statistické nástroje, které zde budou prezentovány, mohou být aplikovány ve všech jmenovaných oblastech, pozornost bude zaměřena převážně na oblast medicíny. Tyto speciální matematické metody jsou označovány názvem *analýza přežívání*. Počátky analýzy přežívání spadají do 17. století, kdy začaly pokusy o zpracování demografických tabulek. Největší rozkvět této metody ovšem nastal ve 20. století po skončení 2. světové války. V posledním desetiletí se aplikace statistických metod analýzy dat rozšířily za hranice biomedicíny do oblastí jako je třeba sociologie, kriminologie a marketing.

Hlavním znakem datových souborů, které zkoumáme metodami analýzy přežívání, je fakt, že kromě *konečných* dat mohou obsahovat i data *cenzorovaná*. Cenzorovaná data vznikají, pokud v době sledování nedojde u daného jedince k výskytu události. Abychom se vyhnuli případným nedorozuměním, stanovíme si jasně pojem události. Událostí se rozumí výskyt libovolného stanoveného jevu, chápaného ze subjektivního pohledu pozorovatele pozitivně, negativně nebo i neutrálně. I když je možno se v konkrétních případech setkat s různým pojmenováním pro událost, obecně bývá označována jako *selhání*.

Pro názornost a lepší pochopení jsou uvedeny některé příklady pro analýzu přežívání z biomedicíny a dalších oblastí, události jsou zvýrazněny proloženým typem písma:

## Biomedicína

- doba přežití - do *úmrtí* (v měsících) pacientů s leukémií, kterým byla transplantována kostní dřeň
- délka remise (v měsících) pacientů s akutní leukémií (do výskytu *relapsu* nebo *úmrtí*) od klinického pokusu podání léku 6-mercaptopurine
- doba do *výskytu infekce* (v měsících) u pacientů s renální insuficiencí při použití různých katetrizačních postupů
- doba přežití - do *úmrtí* (v měsících) pacientek s rakovinou prsu s různou imunohistochemickou odpovědí
- doba působení analgezie po operaci ledvin u kojenců a batolat do *podání* opiátu z důvodu pozorování projevů bolesti



- doba (ve dnech) do *úmrtí* nebo *relapsu* pro pacienty s transplantací kostní dřeně s Hodgkinovým a non-Hodgkinovým lymfomem
- doba (v měsících) do *reinfekce* pacientů se sexuálně přenosnými nemocemi
- doba (v letech) do *první hospitalizace* kvůli pneumonii u malých dětí s pneumonií (kojených a nekojených)
- doba (v letech) do *prvního užití marihuany* u středoškolské mládeže

### **Sociologie**

- doba (v měsících) do *znovuzatčení* recidivistů
- délka trvání manželství (v letech) do *rozvodu*
- doba (v letech) do *ukončení* prvního pracovního poměru u zdravotních sester

### **Inženýrství**

- doba (v hodinách) životnosti součástek průmyslového zařízení do *selhání*

### **Ekonomie**

- délka (v měsících) nezaměstnanosti uchazečů o práci do *uzavření pracovního poměru*
- doba (v měsících) registrace nemovitosti u realitní kanceláře do jejího *prodeje*
- doba (v měsících) od získání půjčky do jejího *plného splacení*

### **Marketing**

- délka (v letech) předplatného časopisu nebo novin do jejího *zrušení*

### **Kriminologie**

- doba (v měsících) do *uzavření případu*
- doba (v měsících) do *uzavření soudního procesu*

### **Vojenský výzkum**

- doba (v minutách) do *zasažení* určitého pohyblivého cíle (např. tank, letadlo, voják) na bojišti

## **Cíle analýzy přežívání jsou**

- odhadnout a interpretovat funkci přežití a/nebo rizikovou funkci
- porovnat funkce přežití a/nebo rizikové funkce
- odhadnout vztah mezi sledovanými vysvětlujícími proměnnými a dobou přežití.

## **1.1 Analýza přežívání**

Obecně je analýza přežívání soubor statistických metod pro analýzu dat, pro něž je *výslednou proměnnou (outcome variable)*, která nás zajímá, *čas do výskytu nějakého jevu*.

*Časem (time) nebo dobou přežití (survival time)* rozumíme počet časových jednotek (let, měsíců, týdnů, dní apod.) od začátku sledování jedince až do výskytu jevu.

*Jevem (event)* myslíme úmrtí, incidenci onemocnění, relaps z remise, selhání transplantovaného orgánu, uzdravení (návrat do práce) nebo jiné určité události, které se mohou vyskytnout u jedince.

Ačkoliv v jedné analýze přežívání může být sledováno více jevů než pouze jeden, budeme se pro názornost metody zabývat pouze situacemi s právě jednou událostí.

## **1.2 Cenzorování**

Většina studií zabývajících se analýzou přežívání je v praxi ukončena dříve než u všech subjektů (pacientů) nastane sledovaná událost. Tuto situaci nazýváme v analýze přežívání *cenzorováním*. Cenzorování nastává tehdy, pokud daný subjekt po jistou dobu sledujeme, ale sledování je ukončeno dříve, než u tohoto subjektu nastala sledovaná událost.

### **Cenzorování může nastat ze tří důvodů:**

1. sledovaná událost u subjektu nenastane do konce studie
2. subjekt je v průběhu sledování ze studie vyřazen nebo sám dobrovolně odstoupí
3. subjekt je ztracen pro pozorování ještě před ukončením studie (např. zemře – pokud smrt není sledovanou událostí).

## Typy cenzorování

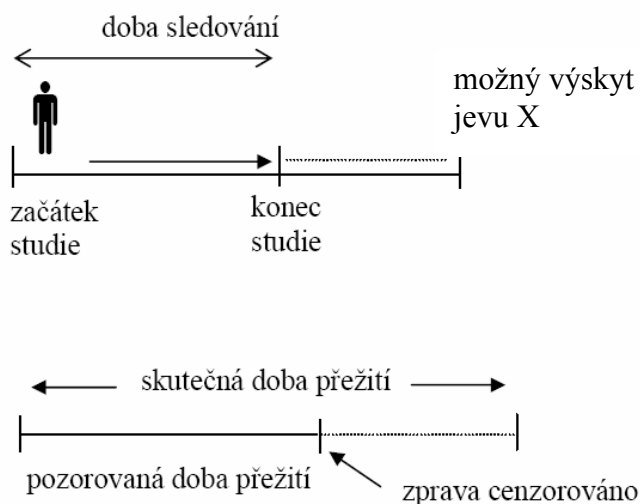
1. *cenzorování zprava (right censoring)* – je nejčastější situací. K nekompletnosti časového údaje dojde na pravé straně sledovaného intervalu (viz. obr. 1.1). Víme pouze, že skutečná doba do události je větší než doba pozorovaná nebo událost u jedince nikdy nenastane.
2. *cenzorování zleva (left censoring)* – k neúplnosti informace o čase selhání dochází na levé straně sledovaného časového intervalu, např. u HIV infikované osoby sledování začíná po prvním pozitivním testu, avšak přesný čas napadení virem neznáme (viz obr. 1.2).
3. *intervalové cenzorování (interval censoring)* – je kombinací cenzorování zleva a zprava, přesnou dobu sledování neznáme, ale víme, že sledovaný jev nastane někdy během známého časového intervalu, např. sledování projevů příznaků nemoci u pacientů při pravidelných půlročních lékařských prohlídkách.

Literatura rovněž uvádí jiné způsoby klasifikace cenzorování, např.:

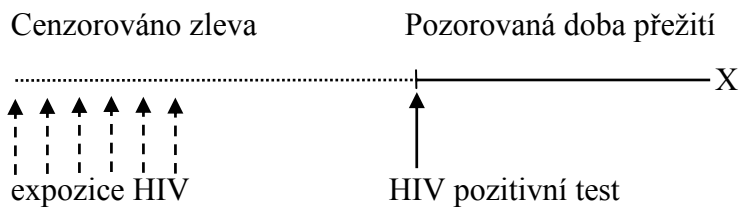
*Cenzorování typu I*, kde každý sledovaný jedinec je pozorován po stejnou dobu; pokud se událost v tomto konstantním intervalu nevyskytne, dojde k cenzorování zprava.

*Cenzorování typu II* – za pozorovanou událost se zde stanoví n-tý výskyt určeného jevu (události).

Obr. 1.1 Cenzorování zprava



Obr. 1.2 Cenzorování zleva



V dalších úvahách se budeme zabývat výlučně daty cenzorovanými zprava.

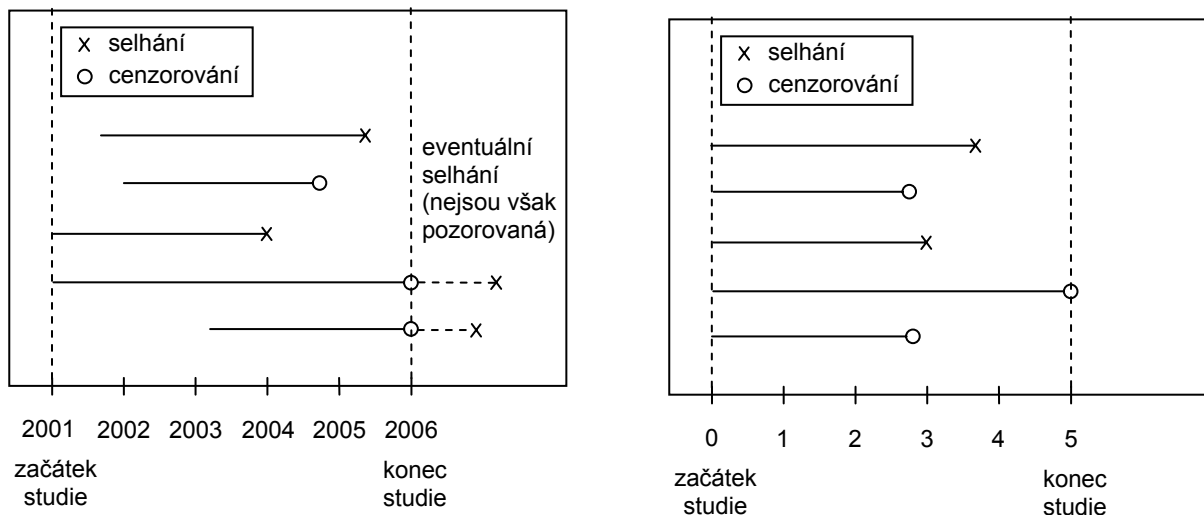
V tabulce 1.1 je ukázáno, jakým způsobem se zadávají data pro analýzu přežívání.

Tab. 1.1 Příklad dat pro analýzu přežití:

Pacient	Doba přežití (měsíce)	Status (1 – jev nastal, 0 – jev nenastal)
A	6	0
B	7,5	1
C	4	1
D	3,5	0
E	5	1
F	8	0

Obr. 1.3 Přežívání pěti subjektů, které vstoupily do studie v různých časových okamžicích.

Vlevo skutečný kalendářní čas, vpravo doba od vstupu do studie.



## Základní pojmy

Nezáporná náhodná veličina  $T$ , která má význam doby uplynulé od zahájení pozorování jedince do výskytu určité, přesně specifikované události, se nazývá *doba přežití*.

$t$  označuje konkrétní hodnotu pro  $T$ .

Náhodná veličina  $\varepsilon = (0, 1)$  označuje fakt, zda došlo k dané události ( $\varepsilon = 1$ ) či nikoliv ( $\varepsilon = 0$ ).

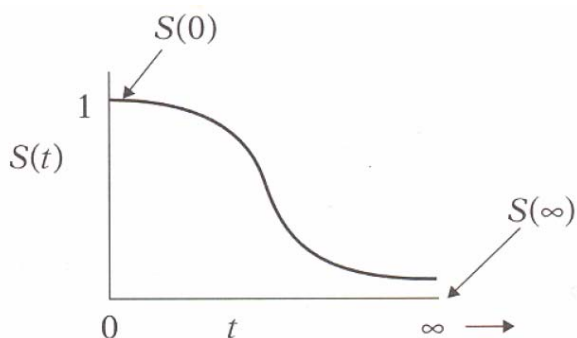
### 1.3 Funkce přežití

Jedním z cílů analýzy přežívání je odhadnout *funkci přežití* (survival function, survivor function)  $S(t)$  a rizikovou funkci  $h(t)$ . Definujme:

Definujme:  $S(t) = P(T > t)$

Funkce přežití udává pravděpodobnost, že osoba žije déle než konkrétně specifikovaný čas  $t$ .

Obr. 1.4 Teoretický tvar funkce přežití  $S(t)$



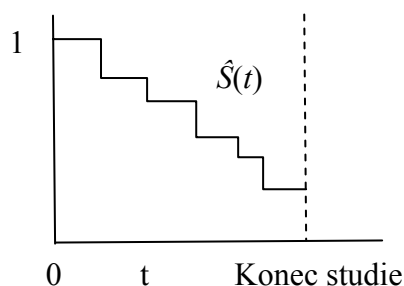
Teoreticky, pro celou populaci, může být funkce přežití zobrazena jako hladká křivka.

Vlastnosti funkce přežití:

- $S(t)$  je nerostoucí funkce
- v čase  $t = 0$  je  $S(t) = S(0) = 1$ , tj. na začátku studie u žádného subjektu nenastal sledovaný jev, proto je pravděpodobnost přežití času  $t = 0$  rovna 1
- v čase  $t = \infty$  je  $S(t) = S(\infty) = 0$ , tj. teoreticky, překročí-li doba sledování určitý limit, žádná osoba nebude přežívat a funkce přežití klesne k nule.

V praxi, použijeme-li konkrétní data, není křivka přežití hladká, ale je spíše *schodovitá* (protože se funkce přežití mění pouze v čase, kdy nastane sledovaný jev).

Obr. 1.5 Odhad tvaru funkce přežití  $\hat{S}(t)$



Protože interval sledování klinické studie není nikdy nekonečný, je možné a v praxi časté, že událost nenastane u všech subjektů a odhadovaná funkce přežití neklesne na konci studie až k nule.

## 1.4 Funkce hazardu

Kromě funkce přežití odhadujeme v analýze přežívání i *rizikovou funkci* neboli *funkci hazardu*  $h(t)$ .

Tato funkce je matematicky definována

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}. \quad (1.1)$$

Funkce hazardu  $h(t)$  udává míru rizika výskytu daného jevu v daném časovém okamžiku. Z definice funkce hazardu vyplývá, že hodnota funkce hazardu v daném čase  $t$  udává pravděpodobnost, že k jevu došlo v intervalu  $(t, t+\Delta t)$ , za podmínky, že k jevu nedošlo před časem  $t$ , přičemž tato pravděpodobnost je vztažena k nekonečně malému jednotkovému intervalu času.

Vlastnosti funkce hazardu

- je nezáporná,  $h(t) \geq 0$  pro všechna  $t$
- není omezená shora (nemá horní hranici)
- v čase 0 nabývá libovolnou hodnotu, totéž platí pro  $t$  jdoucí k nekonečnu.

### Vztah mezi $h(t)$ a $S(t)$

Mezi  $h(t)$  a  $S(t)$  je definován vzájemně jednoznačný vztah. Známe-li funkci přežití  $S(t)$ , můžeme odvodit odpovídající funkci hazardu  $h(t)$  a naopak.

Vztah mezi  $h(t)$  a  $S(t)$  je možno odvodit následujícím způsobem:

Definujme sledovaný jev J: událost nastala v čase T.

Distribuční funkci veličiny J získáme z definice  $F(t) = P(T \leq t)$

Funkce přežití se definuje jako komplementární funkce k distribuční funkci takto  $S(t) = P(T > t) = 1 - P(T \leq t) = 1 - F(t)$ .

Hustotu pravděpodobnosti veličiny J získáme opět z definice derivací distribuční funkce

$$f(t) = \frac{dF(t)}{dt} = \frac{d(1 - S(t))}{dt} = -\frac{dS(t)}{dt} = -S'(t),$$

$$\text{tedy } S(t) = -\int f(t)dt. \quad (1.2)$$

Ze vztahu je rovněž vidět vyjádření hustoty pravděpodobnosti pomocí funkce přežití.

Funkce hazardu se definuje pomocí podmíněné pravděpodobnosti jevu J takto

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \quad (1.3)$$

Nechť

$$\text{jev A: } t \leq T < t + \Delta t \quad (1.4)$$

$$\text{jev B: } T \geq t \quad (1.5)$$

Podle definice podmíněné pravděpodobnosti platí

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (1.6)$$

Dosadíme do rovnice (1.6) výrazy (1.4) a (1.5):

$$P(t \leq T < t + \Delta t | T \geq t) = \frac{P((t \leq T < t + \Delta t) \wedge (T \geq t))}{P(T \geq t)} = \frac{\int_t^{t+\Delta t} f(t)dt}{\int_t^{\infty} f(t)dt}. \quad (1.7)$$

Využijme vztahu (1.2) mezi hustotou pravděpodobnosti a distribuční funkcí a s užitím Newtonovy-Leibnizovy věty dostáváme:

$$\frac{\int_t^{t+\Delta t} f(t)dt}{\int_t^{\infty} f(t)dt} = \frac{[-S(t)]_t^{t+\Delta t}}{[-S(t)]_t^{\infty}} = \frac{-S(t + \Delta t) + S(t)}{-S(\infty) + S(t)} = -\frac{S(t + \Delta t) - S(t)}{S(t)}. \quad (1.8)$$

Na závěr dosadíme upravený výraz (1.7) do definice funkce hazardu (1.1) a upravme limitu, čímž získáváme hledaný vztah:

$$\begin{aligned}
 h(t) &= \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} = - \lim_{\Delta t \rightarrow 0} \frac{S(t + \Delta t) - S(t)}{\Delta t \cdot S(t)} = - \frac{1}{S(t)} \cdot \lim_{\Delta t \rightarrow 0} \frac{S(t + \Delta t) - S(t)}{\Delta t} = \\
 &= - \frac{S'(t)}{S(t)},
 \end{aligned}$$

tedy:

$$h(t) = - \frac{\frac{dS(t)}{dt}}{S(t)} \quad (1.9)$$

a

$$S(t) = e^{-\int_0^t h(u) du} . \quad (1.10)$$

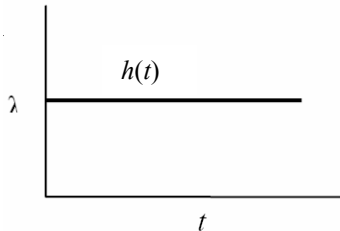


### Příklady funkce hazardu:

**Exponenciální rozdělení**  $\lambda > 0, t \geq 0$ :

Funkce hazardu  $h(t) = \lambda$  (konstatní hazard)

Pro zdravé osoby je riziko onemocnění stejně velké, rovno konstantě.



Funkce přežití  $S(t) = e^{-\lambda t}$

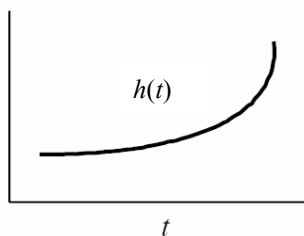
Hustota rozdělení  $f(t) = \lambda e^{-\lambda t}$

Průměr  $E(X) = \frac{1}{\lambda}$

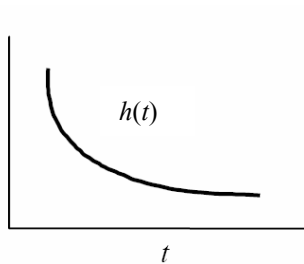
**Weibullovo rozdělení**  $\alpha, \lambda > 0, t \geq 0$ :

Funkce hazardu  $h(t) = \alpha \lambda t^{\alpha-1}$

Rostoucí funkce hazardu pro leukemické pacienty ( $\alpha > 1$ )



Klesající funkce hazardu pro pacienty zotavující se z operace ( $\alpha < 1$ ):



Funkce přežití  $S(t) = e^{-\lambda t^\alpha}$

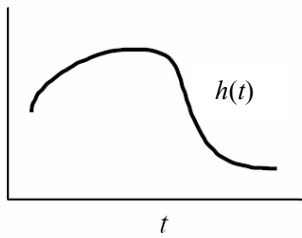
Hustota rozdělení  $f(t) = \alpha \lambda t^{\alpha-1} e^{-\lambda t^\alpha}$

$$\text{Průměr } E(X) = \frac{\Gamma(1 + \frac{1}{\alpha})}{\lambda^{\frac{1}{\alpha}}}$$

**Logaritmicko-normální rozdělení**  $\sigma > 0, t \geq 0$ :

$$\text{Funkce hazardu } h(t) = \frac{f(t)}{S(t)}$$

Funkce hazardu pro pacienty s tuberkulózou:



$$\text{Funkce přežití } S(t) = 1 - \Phi\left[\frac{\ln t - \mu}{\sigma}\right]$$

$$\text{Hustota rozdělení } f(t) = \frac{e^{-\left[\frac{1}{2}\left(\frac{\ln t - \mu}{\sigma}\right)^2\right]}}{t(2\pi)^{\frac{1}{2}}\sigma}$$

$$\text{Průměr } E(X) = e^{(\mu + 0,5\sigma^2)}$$

## 2 Odhady distribuce přežívání

### 2.1 Kaplan-Meierův odhad distribuce přežívání

Tato metoda poskytuje odhad funkce přežití v každém okamžiku, ve kterém došlo k alespoň jedné sledované události.

Kaplan-Meierův odhad funkce přežití v čase  $t_j$  je definován následovně:

$$\hat{S}(t_j) = \prod_{i=1}^j P(T > t_i | T \geq t_i) = \hat{S}(t_{j-1}) P(T > t_{(j)} | T \geq t_{(j)}). \quad (2.1)$$

Tento vztah lze odvodit následujícím způsobem:

Položme:

$$A = \text{„}T \geq t_{(j)}\text{“} \sim \text{„}T > t_{(j-1)}\text{“}$$

jev A nastane, pokud se sledovaná událost projeví v následujících časech  $t_{(j+k)}$ ,  $k \in \{1, 2, 3, \dots\}$ , ekvivalentně se dá vyjádřit, že jev A nastane pokud se sledovaná událost vyskytne později než v čase  $t_{(j-1)}$ . Přičemž je zřejmé, že událost může nastat jen ve známých diskrétních časech.

Dále položme

$$B = \text{„}T > t_{(j)}\text{“}$$

jev B nastane, pokud se sledovaná událost přihodí v libovolném diskrétním čase větším než  $t_{(j)}$ , tedy v časech  $t_{(j+1)}$ ,  $t_{(j+2)}$  atd.

Je zřejmé, že  $A \cap B = B$  a tedy  $P(B) = P(A \cap B)$ .

Podle definice podmíněné pravděpodobnosti  $P(A|B) = \frac{P(A \cap B)}{P(B)}$  lze psát

$$P(B) = P(A \cap B) = P(A) \cdot P(A | B). \quad (2.2)$$

Z definice funkce přežití plyne  $P(B) = P(T > t_{(j)}) = S(t_{(j)})$  a

$$P(A) = P(T \geq t_{(j)}) = P(T > t_{(j-1)}) = S(t_{(j-1)}).$$

Dosazením do vztahu (2.2) dostáváme

$$S(t_{(j)}) = S(t_{(j-1)}) \cdot P(T > t_{(j)} | T \geq t_{(j)}).$$

Tato podmíněná pravděpodobnost se v praxi určuje jako relativní četnost, kde v čitateli je počet subjektů, u kterých událost nenastala ani v čase  $t_{(j)}$  a ve jmenovateli počet subjektů, u nichž událost nenastala ani v čase  $t_{(j-1)}$ .

Tuto teorii si ozřejmíme na následujícím příkladu klinických dat.

### **Příklad - studie z Kliniky plicních nemocí a tuberkulózy FN Olomouc**

U 20 pacientů po radikální operaci pro bronchogenní karcinom léčených adjuvantní chemoterapií byla zaznamenána doba do úmrtí v měsících.

Necenzorovaná data: 14, 17, 21, 21, 27, 28, 50, 99.

Cenzorovaná data: 27+, 49+, 56+, 79+, 85+, 90+, 92+, 93+, 98+, 104+, 109+, 121+.

Sledovaný jev (event) = úmrtí.

Úkolem je odhadnout pravděpodobnosti přežití ve sledovaných časech a sestavit funkce přežití a hazardu.

Postup výpočtu odhadu pravděpodobností je zřejmý z tabulky 2.1, kde se aplikuje vzorec (2.1).

Tab. 2.1. Výpočet pravděpodobnosti přežití

<b>Čas (doba do úmrtí) v měsících</b>	<b>Status Úmrtí = 1</b>	<b>Relativní podíl přežívajících pacientů v čase <math>t_j</math></b>	<b>Odhad pravděpodobnosti přežití <math>\hat{S}(t_j)</math></b>
14	1	19/20	$1 \times 19/20 = \mathbf{0,95}$
17	1	18/19	$0,95 \times 18/19 = \mathbf{0,90}$
21	1		
21	1	16/18	$0,90 \times 16/18 = \mathbf{0,8}$
27	1	15/16	$0,8 \times 15/16 = \mathbf{0,75}$
27	0		
28	1	13/14	$0,75 \times 13/14 = \mathbf{0,696}$
49	0		
50	1	11/12	$0,696 \times 11/12 = \mathbf{0,638}$
56	0		
79	0		
85	0		
90	0		
92	0		
93	0		
98	0		
99	1	3/4	$0,638 \times 3/4 = \mathbf{0,479}$
104	0		
109	0		
121	0		

Obdobné výsledky dostaneme při užití statistického software SPSS verze 15 viz. tabulka 2.2.

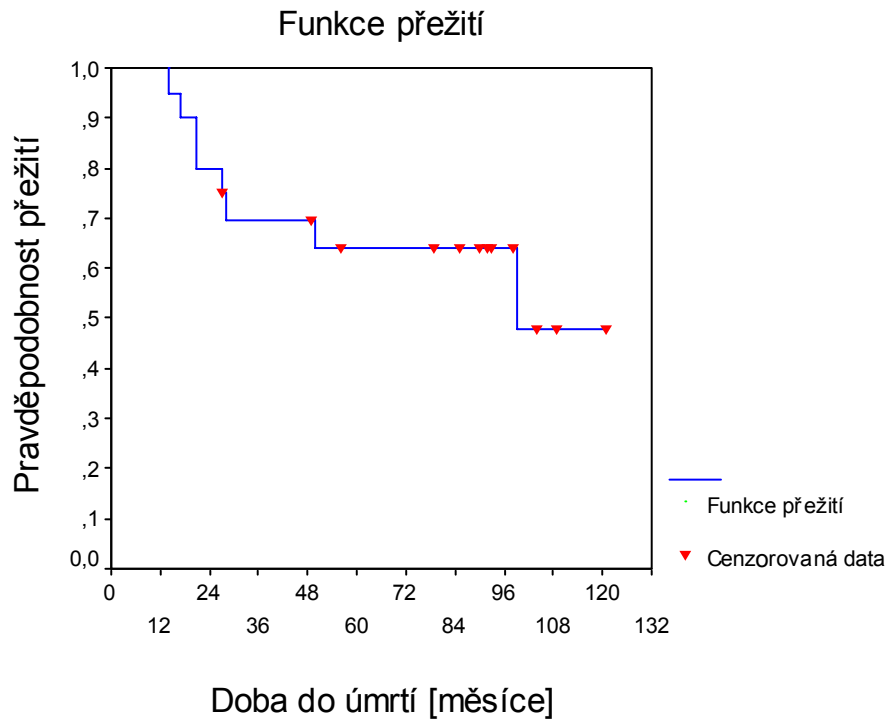
Tab. 2.2 Počítačový výstup Kaplan-Meierovy analýzy přežívání ze statistického programu SPSS.

Survival Table							
skup	Time	Status	Cumulative Proportion Surviving at the Time		N of Cumulative Events	N of Remaining Cases	
			Estimate	Std. Error			
1	1	14,000	1	,950	,049	1	19
	2	17,000	1	,900	,067	2	18
	3	21,000	1	.	.	3	17
	4	21,000	1	,800	,089	4	16
	5	27,000	1	,750	,097	5	15
	6	27,000	0	.	.	5	14
	7	28,000	1	,696	,104	6	13
	8	49,000	0	.	.	6	12
	9	50,000	1	,638	,110	7	11
	10	56,000	0	.	.	7	10
	11	79,000	0	.	.	7	9
	12	85,000	0	.	.	7	8
	13	90,000	0	.	.	7	7
	14	92,000	0	.	.	7	6
	15	93,000	0	.	.	7	5
	16	98,000	0	.	.	7	4
	17	99,000	1	,479	,161	8	3
	18	104,000	0	.	.	8	2
	19	109,000	0	.	.	8	1
	20	121,000	0	.	.	8	0

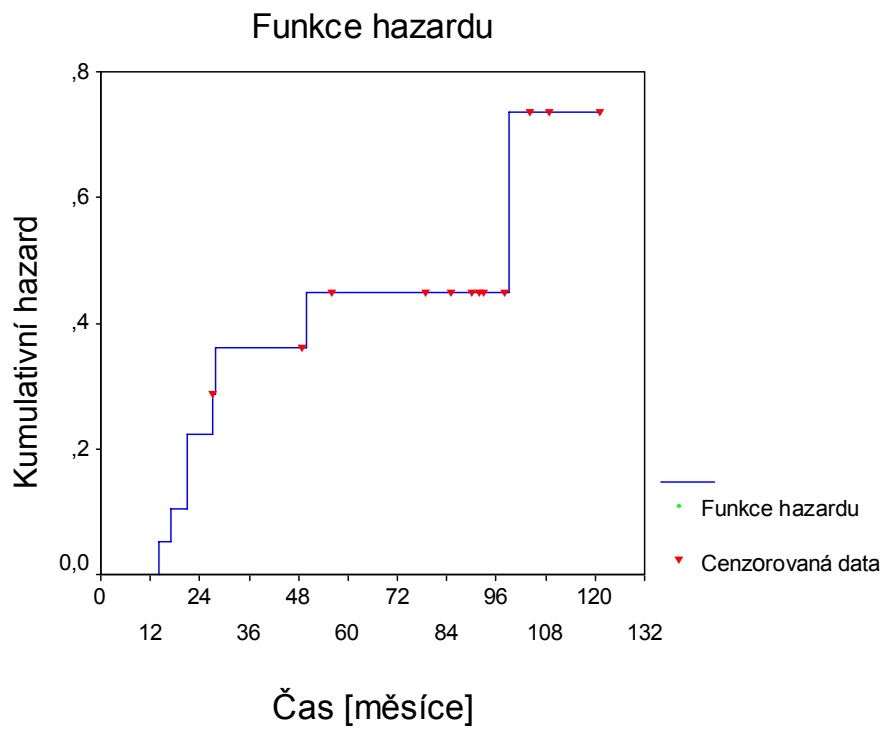
V prvním sloupci jsou uvedena čísla pacientů, ve druhém (Time) čas sledování v měsících, ve sloupci s názvem Status je zaznamenáno, zda pacient zemřel (1) nebo ne (0), další dva sloupce udávají odhad pravděpodobnosti přežívání a chybu tohoto odhadu (Cumulative Proportion Surviving at the Time, Estimate, Std. Error), předposlední sloupec zaznamenává kumulativní počet událostí (úmrtí) a poslední počet případů, které zůstávají ve studii.

Na obrázcích 2.1 a 2.2 je znázorněn průběh funkce přežití a funkce hazardu.

Obr. 2.1



Obr. 2.2



### 3 Porovnání dvou nebo více křivek přežití

Běžným problémem v klinických studiích je srovnání dvou nebo více typů léčení z hlediska jejich schopnosti prodloužit život pacientů nebo udržet pacienty v remisi. Vystává tedy problém porovnat dvě nebo více křivek přežití. Pro toto srovnání existuje několik statistických testů:

Log-rank test, obecný Wilcoxonův test (nazývaný také Breslowův), Tarone-Ware test, Peto-Peto test, Fleming-Harringtonův test. Tyto testy nám odpovídají na otázku, zda jsou Kaplan-Meierovy křivky statisticky ekvivalentní. Přiblížíme si principy nejpoužívanějších dvou testů – log rank testu a Wilcoxonova testu.

#### 3.1 Log-rank test pro dvě skupiny

Log-rank test je v podstatě chí-kvadrát test pro velký výběr. Log-rank statistika, stejně jako jiné statistiky používané v chí-kvadrát testech, porovnává pozorovaný počet jevů (událostí) s očekávaným za platnosti nulové hypotézy.

Testová statistika pro dvě skupiny má tvar:

$$\text{Log-rank statistika} = \frac{(O_i - E_i)^2}{\text{Var}(O_i - E_i)}, \quad (3.1)$$

$$\text{kde } O_i - E_i = \sum_{j=1}^n (m_{ij} - e_{ij}), \quad i = 1, 2 \quad (3.2)$$

$m_{ij}$  je pozorovaný počet jevů v souboru  $i$  v čase  $j$ ,

$e_{ij}$  je očekávaný počet jevů v souboru  $i$  v čase  $j$ ,

$$e_{1j} = \left( \frac{n_{1j}}{n_{1j} + n_{2j}} \right) (m_{1j} + m_{2j}), \quad \text{kde } m_{1j}, m_{2j} \text{ je počet událostí v souboru 1 a 2 v čase } j. \quad (3.3)$$

↑                    ↑  
Podíl                Počet  
v rizikové        událostí  
skupině            u obou  
                         skupin

$n_{ij}$  – počet subjektů v  $i$ -té skupině a v  $j$ -tém čase selhání

$$e_{2j} = \left( \frac{n_{2j}}{n_{1j} + n_{2j}} \right) \cdot (m_{1j} + m_{2j}) \quad (3.4)$$

$$Var(O_i - E_i) = \sum_j \frac{n_{1j}n_{2j}(m_{1j} + m_{2j})(n_{1j} + n_{2j} - m_{1j} - m_{2j})}{(n_{1j} + n_{2j})^2(n_{1j} + n_{2j} - 1)}, i = 1, 2 \quad (3.5)$$

Hypotéza  $H_0$ : Neexistuje rozdíl mezi křivkami přežití.

Statistika log-rank má  $\chi^2$  rozdělení s 1 stupněm volnosti.

### **Příklad - studie z Kliniky plicních nemocí a tuberkulózy FN Olomouc, pokračování**

Studie si kladla za cíl posoudit přežívání 2 skupin pacientů po radikální operaci pro bronchogenní karcinom léčených adjuvantní chemoterapií v závislosti na typu léčby, kterou podstoupili.

1. skupina - pacienti, kterým byla aplikována cytostatika podle MTT testu (byla zjišťována citlivost in vitro nádorových buněk na testovaná cytostatika):

20 pacientů, doba do úmrtí v měsících.

Necenzorovaná data: 14, 17, 21, 21, 27, 28, 50, 99.

Cenzorovaná data: 27+, 49+, 56+, 79+, 85+, 90+, 92+, 93+, 98+, 104+, 109+, 121+.

Výpočet pravděpodobnosti přežití je uveden v tabulce 2.1.

2. skupina - pacienti, kterým byly aplikována cytostatika bez MTT testu:

24 pacientů, doba do úmrtí v měsících.

Necenzorovaná data: 12, 14, 20, 21, 23, 27, 31, 32, 33, 42, 45, 49, 59, 69.

Cenzorovaná data: 32+, 37+, 51+, 55+, 55+, 60+, 90+, 91+, 92+, 99+.

Sledovaný jev = úmrtí.

Výpočet pravděpodobnosti přežití je uveden v tabulce 3.1.



Tab. 3.1 Výpočet pravděpodobnosti přežití

Pacienti léčeni léky bez MTT testu:

Čas (doba do úmrtí) v měsících	Status Úmrtí = 1	Relativní podíl přežívajících pacientů v čase $t_j$	Odhad pravděpodobnosti přežití $\hat{S}(t_j)$
12	1	23/24	$1 \times 23/24 = 0,9583$
14	1	22/23	$0,9583 \times 22/23 = 0,9167$
21	1	21/22	$0,9167 \times 21/22 = 0,875$
23	1	20/21	$0,875 \times 20/21 = 0,833$
27	1	19/20	$0,833 \times 19/20 = 0,7917$
29	1	18/19	$0,7917 \times 18/19 = 0,75$
31	1	17/18	$0,75 \times 17/18 = 0,7083$
32	1	16/17	$0,7083 \times 16/17 = 0,6667$
32	0		
33	1	14/15	$0,6667 \times 14/15 = 0,622$
37	0		
42	1	12/13	$0,622 \times 12/13 = 0,5744$
45	1	11/12	$0,5744 \times 11/12 = 0,5265$
49	1	10/11	$0,5265 \times 10/11 = 0,4786$
51	0		
55	0		
55	0		
59	1	6/7	$0,4786 \times 6/7 = 0,4103$
60	0		
69	1	4/5	$0,4103 \times 4/5 = 0,3282$
90	0		
91	0		
92	0		
99	0		

Abychom mohli porovnat přežívání 2 skupin pacientů po radikální operaci pro bronchogenní karcinom léčených adjuvantní chemoterapií v závislosti na typu léčby, kterou podstoupili, je nutné spočítat očekávané počty událostí v jednotlivých časových krocích. Tento výpočet dle vzorce (3.3) je proveden v tabulce 3.2.

Tab. 3.2 Výpočet očekávaných počtů událostí

j	$t_{(j)}$	Pozorovaný počet událostí		Počet pacientů v rizikové skupině		Očekávaný počet událostí		Pozorovaný – očekávaný počet událostí	
		$m_{1j}$	$m_{2j}$	$n_{1j}$	$n_{2j}$	$e_{1j}$	$e_{2j}$	$m_{1j} - e_{1j}$	$m_{2j} - e_{2j}$
1	12	0	1	20	24	(20/44).1=0,45	(24/44).1=0,55	-0,45	0,45
2	14	1	1	20	23	(20/43).2=0,93	(23/43).2=1,07	0,07	-0,07
3	17	1	0	19	22	(19/41).1=0,46	(22/41).1=0,54	0,54	-0,54
4	21	2	1	18	22	(18/40).3=1,35	(22/40).3=1,65	0,65	-0,65
5	23	0	1	16	21	(16/37).1=0,43	(21/37).1=0,57	-0,43	0,43
6	27	1	1	16	20	(16/36).2=0,89	(20/36).2=1,11	0,11	-0,11
7	28	1	0	14	19	(14/33).1=0,42	(19/33).1=0,58	0,58	-0,58
8	29	0	1	13	19	(13/32).1=0,40	(19/32).1=0,60	-0,40	0,40
9	31	0	1	13	18	(13/31).1=0,42	(18/31).1=0,58	-0,42	0,42
10	32	0	1	13	17	(13/30).1=0,43	(17/30).1=0,57	-0,43	0,43
11	33	0	1	13	15	(13/28).1=0,46	(15/28).1=0,54	-0,46	0,46
12	42	0	1	13	13	(13/26).1=0,5	(13/26).1=0,5	-0,5	0,5
13	45	0	1	13	12	(13/25).1=0,52	(12/25).1=0,48	-0,52	0,52
14	49	0	1	13	11	(13/24).1=0,54	(11/24).1=0,46	-0,54	0,54
15	50	1	0	12	10	(12/22).1=0,55	(10/22).1=0,45	0,45	-0,45
16	59	0	1	11	10	(11/21).1=0,52	(10/21).1=0,48	-0,52	0,52
17	69	0	1	10	5	(10/15).1=0,67	(5/15).1=0,33	-0,67	0,67
18	99	1	0	4	2	(4/6).1=0,67	(2/6).1=0,33	0,33	-0,33
<b>Součet</b>		<b>8</b>	<b>14</b>			<b>10,61</b>	<b>11,39</b>	<b>-2,61</b>	<b>2,61</b>

Podle vztahů (3.2), (3.5) a (3.1) dostáváme:

$$O_2 - E_2 = \sum_j (m_{ij} - e_{ij}) = 2,61$$

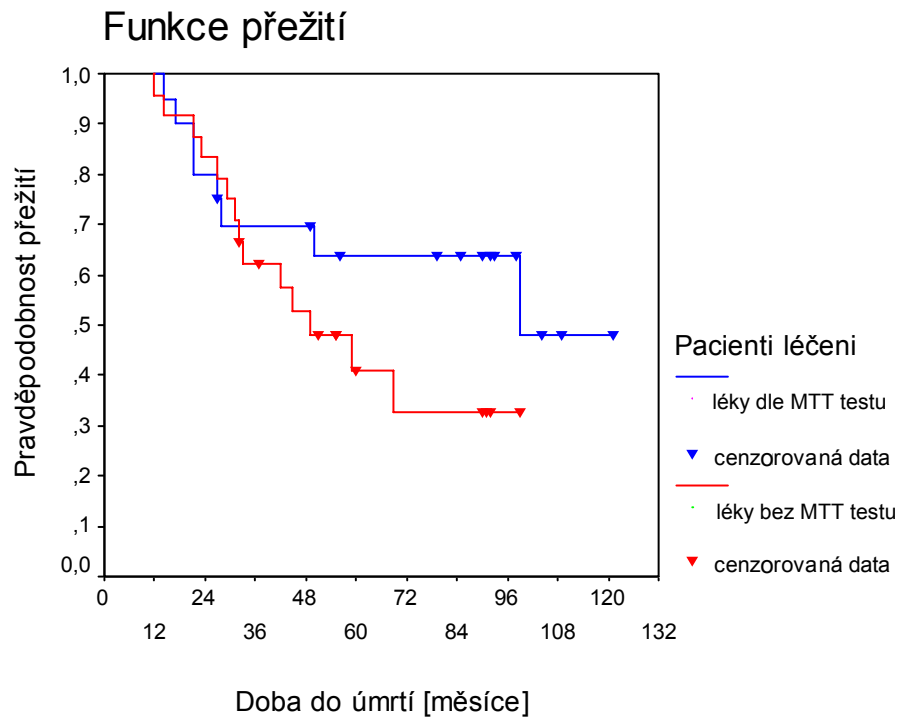
$$Var(O_2 - E_2) = \sum_{j=1}^{18} \frac{n_{1j}n_{2j}(m_{1j} + m_{2j})(n_{1j} + n_{2j} - m_{1j} - m_{2j})}{(n_{1j} + n_{2j})^2(n_{1j} + n_{2j} - 1)} = 12,03$$

$$\text{Log-rank statistika} = \frac{(O_2 - E_2)^2}{Var(O_2 - E_2)} = 0,217.$$

Hodnota log-rank statistiky nepřekročila kvantil  $\chi^2_{1,0,95} = 3,84$ . Nulovou hypotézu tedy nelze zamítnout na 5% hladině statistické významnosti. Křivky přežívání pacientů s rakovinou plic léčených I.a II. typem léčby se statisticky prokazatelně neliší.

Graficky je situace znázorněna na obr. 3.1.

Obr. 3.1 Graf Kaplan-Meierových křivek přežití



### 3.2 Log-rank test pro více skupin

Log-rank test může být také použit k porovnání tří nebo více křivek přežití. Nulová hypotéza předpokládá, že všechny křivky přežití jsou stejné.

Pro  $i = 1, 2, \dots, G$  a  $j = 1, 2, \dots, k$ , kde  $G$  je počet skupin a  $k$  je počet diskretních časových okamžiků, ve kterých došlo k selhání, je

$n_{ij}$  – počet subjektů v  $i$ -té skupině a v  $j$ -tém čase selhání

$m_{ij}$  – pozorovaný počet selhání v  $i$ -té skupině a v  $j$ -tém čase selhání

$e_{ij}$  – očekávaný počet selhání v  $i$ -té skupině a v  $j$ -tém čase selhání

$$e_{ij} = \left( \frac{n_{ij}}{n_{1j} + n_{2j}} \right) (m_{1j} + m_{2j})$$

$$n_j = \sum_{i=1}^G n_{ij}$$

$$m_j = \sum_{i=1}^G m_{ij}$$

$$O_i - E_i = \sum_{j=1}^k (m_{ij} - e_{ij})$$

$$Var(O_i - E_i) = \sum_{j=1}^k \frac{n_{ij}(n_j - n_{ij})m_j(n_j - m_j)}{n_j^2(n_j - 1)}$$

$$Cov(O_i - E_i, O_l - E_l) = \sum_{j=1}^k \frac{-n_{ij}n_{lj}m_j(n_j - m_j)}{n_j^2(n_j - 1)}$$

$$\mathbf{d} = (O_1 - E_1, O_2 - E_2, \dots, O_{G-1} - E_{G-1})'$$

$$\mathbf{V} = ((v_{il}))'$$

kde  $v_{ii} = Var(O_i - E_i)$  a  $v_{il} = Cov(O_i - E_i, O_l - E_l)$  pro  $i = 1, 2, \dots, G-1$ ;  $l = 1, 2, \dots, G-1$ .

Potom log-rank statistika je dána následující formulí:

$$\text{Log-rank statistika} = \mathbf{d}'\mathbf{V}^{-1}\mathbf{d},$$

která má přibližně chí-kvadrát rozdělení s  $G-1$  stupni volnosti za platnosti nulové hypotézy, která předpokládá, že všechny skupiny mají stejnou křivku přežití.

## Příklad

Posouzení přežívání pacientů po radikální operaci pro bronchogenní karcinom v závislosti na 4 typech léčby:

Zpracováno programem SPSS:

**Case Processing Summary**

skup	Total N	N of Events	Censored	
			N	Percent
adjuvance Gemzar	13	5	8	61,5%
vinorelbin	14	9	5	35,7%
taxany	5	2	3	60,0%
neplatinové režimy	11	5	6	54,5%
Overall	43	21	22	51,2%

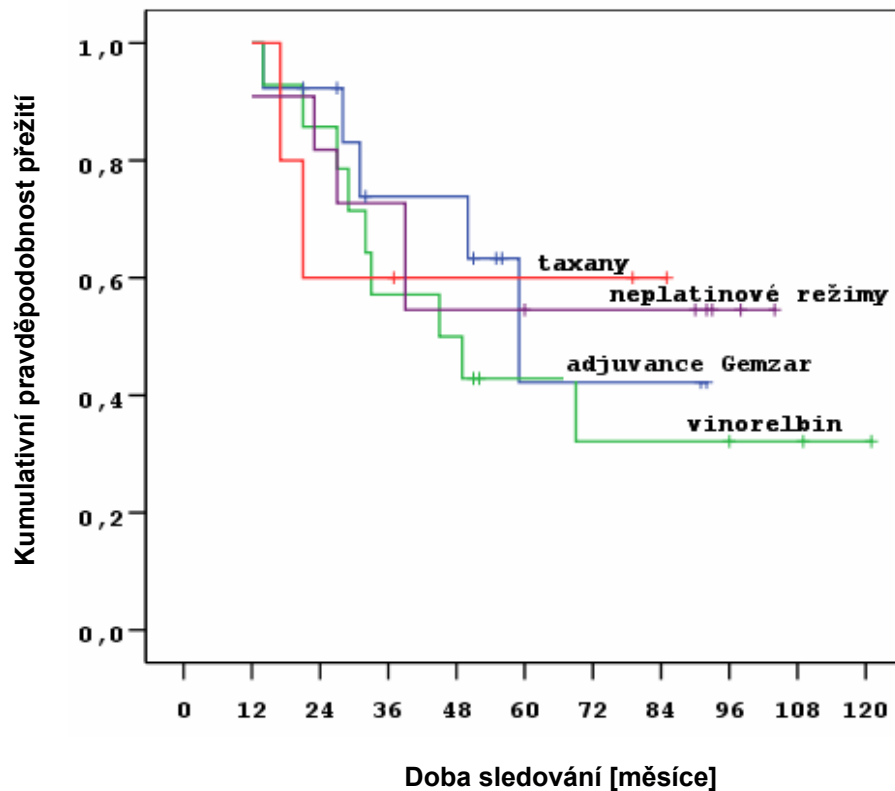
První skupina 13 pacientů užívala léčivo Gemzar, 5 pacientů během studie zemřelo, 8 dat bylo cenzorováno.

Druhá skupina 14 pacientů užívala léčivo Vinorelbin, 9 pacientů během studie zemřelo, 5 dat bylo cenzorováno.

Třetí skupina 5 pacientů užívala k léčbě taxany, 2 pacienti během studie zemřeli, 3 data byla cenzorována.

Čtvrtá skupina 11 pacientů byla léčena neplatinovými léčivy, 5 pacientů během studie zemřelo, 6 dat bylo cenzorováno.

Kaplan-Meierovy křivky přežití pro pacienty s různými léčebnými režimy:



Log-rank test porovnávající distribuci přežívání čtyř skupin pacientů:

**Overall Comparisons**

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	1,002	3	,801

Test of equality of survival distributions for the different levels of skup.

Z dosažené hladiny signifikance log-rank testu  $p = 0,801$  je zřejmé, že distribuce přežívání sledovaných skupin pacientů se statisticky prokazatelně neliší.

### 3.3 Wilcoxonův test pro porovnání dvou křivek přežití

Wilcoxonův test porovnávající distribuci dvou křivek přežití je rozšířením Wilcoxonova testu pro vzorky dat zprava cenzorovaných.

Označme  $t$  čas, který uplynul od zahájení studie do selhání (pokud k němu u daného pacienta došlo) nebo do cenzorování,  $n_1, n_2$  počty subjektů v první a druhé skupině. Obecně u  $n_i - r_i$  jedinců došlo k selhání a  $r_i$  jedinců bylo cenzorováno v čase  $t$  ( $i = 1, 2$ ).

Alling navrhl modifikovaný Wilcoxonův test pro cenzorovaná pozorování [1].

#### 3.3.1 $W$ statistika

Předpokládejme, že  $n_1, n_2$  je počet jedinců náhodně rozdělených do dvou skupin, první skupina je podrobena léčbě A, druhá léčbě B.

Bylo zjištěno:

$$\left. \begin{array}{ll} x_1', x_2', \dots, x_{r_1}', & r_1 \text{ cenzorovaných dat} \\ x_{r_1+1}, x_{r_1+2}, \dots, x_{n_1}, & n_1 - r_1 \text{ selhání (událostí)} \end{array} \right\} \text{ léčba A}$$

$$\left. \begin{array}{ll} y_1', y_2', \dots, y_{r_2}', & r_2 \text{ cenzorovaných dat} \\ y_{r_2+1}, y_{r_2+2}, \dots, y_{n_2}, & n_2 - r_2 \text{ selhání (událostí)} \end{array} \right\} \text{ léčba B}$$

kde  $x_i, y_j$  jsou hodnoty času do výskytu události a  $x_i', y_j'$  jsou hodnoty času do cenzorování (všechny hodnoty byly měřeny od vstupu jednotlivých subjektů do studie).

$F_1(x), F_2(x)$  jsou kumulativní distribuční funkce, které mohou být diskrétní nebo spojité.

Nulová hypotéza je stanovena následovně:

$$H_0: F_1(t) = F_2(t) \quad (t \leq T) \text{ (léčby A i B jsou stejně účinné).}$$

Alternativní hypotézy jsou buď:

$$H_1: F_1(t) < F_2(t) \quad (t \leq T) \text{ (léčba A je účinnější než léčba B)}$$

nebo oboustranná alternativa

$$H_2: F_1(t) \neq F_2(t) \quad (t \leq T) \text{ (léčby A i B nejsou stejně účinné).,}$$

kde  $T$  je největší zjištěná doba sledování.

Definujeme:

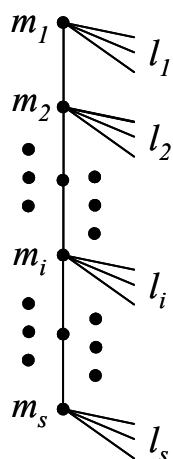
$$U_{ij} = \begin{cases} -1 & \text{jestliže } x_i < y_j \text{ nebo } x_i \leq y_j' \\ 0 & \text{jestliže } x_i = y_j \text{ nebo } (x_i' = y_j') \text{ nebo } x_i' < y_j \text{ nebo } y_j' < x_i \\ +1 & \text{jestliže } x_i > y_j \text{ nebo } x_i' \geq y_j \end{cases} \quad (3.6)$$

a je počítána statistika  $W = \sum_{i,j} U_{ij}$ , kde  $i \in \langle 1, n_1 \rangle$ ,  $j \in \langle 1, n_2 \rangle$  a kde se sčítá přes všechny možné kombinace  $n_1 n_2$ .

Ve statistice  $W$  se odrážejí všechny situace porovnání dat první a druhé skupiny (možné kombinace jsou tyto – selhání ve skupině A vs. selhání ve skupině B, selhání ve skupině A vs. cenzorování ve skupině B, cenzorování ve skupině A vs. selhání ve skupině B, cenzorování ve skupině A vs. cenzorování ve skupině B). V kombinaci, kdy ve skupině A došlo k selhání dříve než k selhání nebo cenzorování ve skupině B, pak je příspěvek k výsledné statistice -1. V kombinaci, kdy ve skupině A došlo k selhání později než k selhání ve skupině B nebo ve skupině A došlo k přežití delšímu než selhání ve skupině B, pak je příspěvek k výsledné statistice +1.

### 3.3.2 Průměr a rozptyl statistiky $W$

Mějme  $n_1 n_2$  pozorování, které mohou být uspořádány podle následujícího vzoru:



Toto uspořádání bylo získáno následujícím postupem:

- veškerá pozorování (konečná i cenzorovaná data) byla seřazena vzestupně podle doby selhání nebo cenzorování
- byly vytvořeny skupiny selhání se stejnou dobou selhání; každou skupinu charakterizuje počet selhání  $m_i$

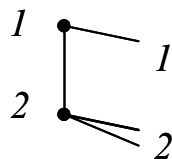


- tím byla získána právě jedna skupina cenzorovaných dat mezi každými dvěma skupinami dat selhání; každé takové skupině určíme číslo  $l_i$  jako počet cenzorovaných dat ve skupině.

Takto definovaný vzor označme písmenem  $P$ .

Například, zaznamenejme tímto schématem následující soubor dat představující přežívání v týdnech: konečná data 8, 10, 10, cenzorovaná data 8+, 11+, 14+.

Schema  $P$ :



Předpokládejme, že  $H_0$  platí a průměrná doba přežívání je stejná v obou skupinách. Uvažujme  $n_1 + n_2$  navzájem odlišitelných jednotlivých pozorování v obou souborech, pak  $\binom{n_1 + n_2}{n_1}$  je počet všech možných uspořádání těchto pozorování do dvou souborů o velikostech  $n_1$  a  $n_2$ .

Uvažujme podmíněný průměr a rozptyl proměnné  $W$  za platnosti nulové hypotézy  $H_0$ . Tento průměr označme  $E(W|P, H_0)$  a rozptyl  $\text{var}(W|P, H_0)$ , kde  $P$  je výše definovaný vzor uspořádání pozorování. Libovolné z možných  $\binom{n_1 + n_2}{n_1}$  uspořádání dat do obou souborů vede na stejný vzor  $P$ .

Vzhledem k symetrii je zřejmé, že

$$E(W|P, H_0) = 0.$$

Výpočet rozptylu je dán vztahem (odvození viz [7] ):

$$\text{var}(W|P, H_0) = \frac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 1)} \left[ \sum_{i=1}^s m_i M_{i-1} (M_{i-1} + 1) + \sum_{i=1}^s l_i M_i (M_i + 1) + \sum_{i=1}^s m_i (n_1 + n_2 - M_i - L_{i-1}) + (n_1 + n_2 - 3M_{i-1} - m_i - L_{i-1} - 1) \right], \quad (3.7)$$

$$\text{kde } M_j = \sum_{i=1}^j m_i, \quad M_0 = 0,$$

$$L_j = \sum_{i=1}^j l_i, \quad L_0 = 0.$$

### 3.3.3 Výpočet statistiky $W$ a $\text{var}(W|P, H_0)$ ve velkých vzorcích

V případě, že počet pozorování  $n_1, n_2$  v obou skupinách je větší než 20, je možno použít zjednodušenou metodu výpočtu statistiky  $W$  a rozptylu  $\text{var}(W)$ . Tato metoda je založena na vytváření  $s$  skupin selhání a cenzorovaných pozorování.

Toto zjednodušení vede k určité ztrátě informací, která je způsobena zanedbáním vzájemného pořadí pozorování uvnitř jednotlivých skupin. Výsledek bude tedy méně přesný než v případě použití výpočtu podle schématu  $P$ .

V rámci metody provedeme uspořádání všech pozorování skupiny aplikující léčbu A podle následujícího vzoru:

Léčba A			
Interval	Počet selhání	Kumulativní počet selhání	Počet cenzorovaných dat
1	$f_{1A}$	$F_{1A}$	$c_{1A}$
.	.	.	.
.	.	.	.
.	.	.	.
$i$	$f_{iA}$	$F_{iA}$	$c_{iA}$
.	.	.	.
.	.	.	.
.	.	.	.
$s$	$f_{sA}$	$F_{sA}$	$c_{sA}$

(3.8)

kde

$f_{iA}$  = počet selhání v intervalu  $i$ ,

$c_{iA}$  = počet cenzorovaných pozorování v intervalu  $i$ ,

$$F_{iA} = \sum_{j=1}^i f_{jA}$$

a obdobně lze vytvořit schéma s položkami  $f_{iB}$ ,  $c_{iB}$  a  $F_{iB}$  pro léčbu B.

Skupiny jsou vytvořeny tak, že se celková doba pozorování skupiny rozdělí na  $s$  navazujících časových intervalů. Následně se určí počty selhání a cenzorování v každém takovém intervalu. Intervaly jsou voleny podobně jako u četnostních distribucí a nemusí být stejné délky.

Aplikací schematu (3.8) na základní vzorce (3.6) obdržíme pro výpočet statistiky  $W$  tento vztah:

$$W = \sum_{i=1}^s [(f_{iA} + c_{iA})F_{i-1,B} - (f_{iB} + c_{iB})F_{i-1,A}], \quad (3.9)$$

kde předpokládáme, že v obou vzorcích jsou použity k výpočtům stejné intervaly.

Pro vstupní proměnné ve výpočtu variability  $\text{var}(W)$  dostáváme tyto vztahy:

$$m_i = f_{iA} + f_{iB}, \quad l_i = c_{i+1,A} + c_{i+1,B}. \quad (3.10)$$

Obě hodnoty  $E(W)$  a  $\text{var}(W)$  počítané touto zjednodušenou metodou mají tendenci být menší než hodnoty získané základní metodou bez seskupování.

### 3.3.4 Příklad

V této kapitole si ukážeme aplikaci Wilcoxonova testu na data, která jsme v kapitole 3.1. použili k objasnění log-rank testu.

Budeme tedy opět posuzovat přežívání 2 skupin pacientů po radikální operaci pro bronchogenní karcinom léčených adjuvantní chemoterapií v závislosti na typu léčby, kterou podstoupili.

1. skupina - pacienti, kterým byl aplikován lék podle MTT testu (byla zjišťována citlivost in vitro nádorových buněk na testovaná cytostatika):

20 pacientů, doba do úmrtí v měsících.

Necenzorovaná data: 14, 17, 21, 21, 27, 28, 50, 99.

Cenzorovaná data: 27+, 49+, 56+, 79+, 85+, 90+, 92+, 93+, 98+, 104+, 109+, 121+.

2. skupina - pacienti, kterým byl aplikován lék bez MTT testu:

24 pacientů, doba do úmrtí v měsících.

Necenzorovaná data: 12, 14, 20, 21, 23, 27, 31, 32, 33, 42, 45, 49, 59, 69.

Cenzorovaná data: 32+, 37+, 51+, 55+, 55+, 60+, 90+, 91+, 92+, 99+.

Protože největší doba sledování je zde 121 měsíců,  $T = 121$

Použijeme-li vztah uvedený v kapitole 3.3.3 pro seskupená data dostáváme schéma:

Interval (měsíce)	Léčba 1				Léčba 2			
	$f_{iA}$	$F_{iA}$	$c_{iA}$	$a_i$	$f_{iB}$	$F_{iB}$	$c_{iB}$	$b_i$
0 – 9	0	0	0	–	0	0	0	–
10 – 19	2	2	0	0	2	2	0	0
20 – 29	4	6	1	10	4	6	0	8
30 – 39	0	6	0	0	3	9	2	30
40 – 49	0	6	1	9	3	12	0	18
50 – 59	1	7	1	24	1	13	3	24
60 – 69	0	7	0	0	1	14	1	14
70 – 79	0	7	1	14	0	14	0	0
80 – 89	0	7	1	14	0	14	0	0
90 – 99	1	8	4	70	0	14	4	28
100 – 109	0	8	2	28	0	14	0	0
110 – 119	0	8	0	0	0	14	0	0
120 – 129	0	8	1	14	0	14	0	0
	<b>8</b>		<b>12</b>	<b>183</b>	<b>14</b>		<b>10</b>	<b>122</b>

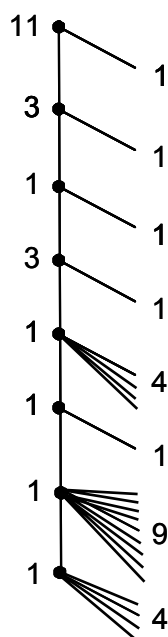
Tedy 
$$W = \sum_{i=1}^{13} (a_i - b_i) = 183 - 122 = 61,$$

kde 
$$a_i = (f_{iA} + c_{iA})F_{i-1,B}, \quad b_i = (f_{iB} + c_{iB})F_{i-1,A}.$$

Rozptyl  $\text{var}(W | P, H_0)$  pro seskupená data spočítáme podle vztahu (3.7), kde

$$m_i = f_{iA} + f_{iB}, \quad l_i = c_{i+1,A} + c_{i+1,B}.$$

Vzor  $P$  z kapitoly 3.3.2 je



Tabulka pro výpočet rozptylu je následující

$i$	$m_i$	$M_i$	$d_i$	$m_i \times d_{i-1}$	$l_i$	$L_{i-1}$	$l_i \times d_i$	$e_i$	$f_i$	$e_i \times f_i$	$m_i \times e_i \times f_i$
1	0	0	0	0	0	0	0	44	43	1892	0
2	4	4	20	0	1	0	20	40	39	1560	6240
3	8	12	156	160	2	1	312	31	22	682	5456
4	3	15	240	468	1	3	240	26	1	26	78
5	3	18	342	720	4	4	1368	22	-9	-198	-594
6	2	20	420	684	1	8	420	16	-21	-336	-672
7	1	21	462	420	1	9	462	14	-27	-378	-378
8	0	21	462	0	1	10	462	13	-30	-390	0
9	0	21	462	0	8	11	3696	12	-31	-372	0
10	1	22	506	462	2	19	1012	3	-40	-120	-120
11	0	22	506	0	0	21	0	1	-44	-44	0
12	0	22	506	0	1	21	506	1	-44	-44	0
13	0	22	506	0							
	<b>22</b>			<b>2914</b>	<b>22</b>		<b>8498</b>				<b>10010</b>

kde  $d_i = M_i(M_i + 1)$ ,  $d_0 = 0$ ,  
 $e_i = n_1 + n_2 - M_i - L_{i-1}$ ,  
 $f_i = n_1 + n_2 - 3M_{i-1} - m_i - L_{i-1} - 1$ .

Pak 
$$\text{var}(W|P, H_0) = \frac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 1)} \left( \sum_{i=1}^{13} m_i d_{i-1} + \sum_{i=1}^{13} l_i d_{i-1} + \sum m_i e_i f_i \right) =$$

$$= \frac{20 * 24}{44 * 43} (2914 + 8498 + 10010) = 5434,757,$$

$$\sqrt{\text{var}(W|P, H_0)} = 73,72$$

Předpokládejme, že testujeme nulovou hypotézu

$$H_0 : F_1(t) = F_2(t), \quad t \leq T \quad \text{proti} \quad \text{oboustranné alternativní hypotéze}$$

$$H_1 : F_1(t) \neq F_2(t), \quad t \leq T,$$

pak

$$Z = \frac{W}{\sqrt{\text{var}(W|P, H_0)}} = \frac{61}{73,72} = 0,8274.$$

Pravděpodobnost takové hodnoty  $Z$  je větší než 0,05, hypotézu  $H_0 : F_1(t) = F_2(t), t \leq T$  nelze zamítnout. Nepodařilo se tedy prokázat, že přežívání pacientů po radikální operaci pro bronchogenní karcinom léčených adjuvantní chemoterapií závisí na typu léčby, kterou podstoupili.

Tento výsledek odpovídá závěru, který jsme učinili po provedení log-rank testu.

## 4 Coxův regresní model

Vícenásobná regresní metoda je konvenční technika vyšetření vztahu mezi dobou přežití a možnými vysvětlujícími proměnnými. Odhalí, nejen které faktory jsou statisticky významné, ale určí i míru, s jakou jednotlivé faktory přežívání ovlivňují.

### 4.1 Coxův model proporcionálního hazardu

*Coxův model poměrného hazardu (Cox proportional hazards model)* byl navržen D. R. Coxem v roce 1972. V české literatuře se také používá název *Coxův model proporcionálního rizika* nebo *Coxova regrese*.

K prokázání vlivu určitých proměnných (faktorů) na přežití a riziko výskytu sledovaného jevu používáme metodu Coxovy regrese. Tato vícenásobná regresní metoda odhalí statisticky významné faktory a určí i míru vlivu jednotlivých faktorů na přežívání.

Vstupní údaje pro Coxovu regresi:

- doba do výskytu sledovaného jevu
- důvod ukončení sledování – kódujeme opět: 1 pokud nastala událost, 0 pro cenzorovaná data
- nezávisle proměnné prediktory – faktory (mohou to být kvantitativní i kvalitativní veličiny).

Coxův model pro  $p$  nezávisle proměnných  $X_1, \dots, X_p$  má tvar:

$$h(t) / h_0(t) = e^{(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)}, \quad (4.1)$$

kde  $h(t)$  je funkce hazardu,  $h_0(t)$  je funkce základního hazardu (tj. očekávané riziko bez vlivu uvažovaných faktorů),  $\beta_1, \dots, \beta_p$  regresní koeficienty.

Veličina  $h(t)/h_0(t)$  se nazývá relativní riziko nebo hazard (HR) a indikuje růst nebo pokles rizika způsobený vlivem faktorů  $X_1, \dots, X_p$ .

Coxova regresní metoda není založena na předpokladu žádného konkrétního rozložení pravděpodobnosti. Avšak předpokládá, že vliv uvažovaných faktorů je v čase konstantní a má aditivní efekt.

Odhad koeficientů Coxovy regresní rovnice je dnes součástí statistických softwarových aplikací. Pomocí složitých matematických metod jsou odhadnuty koeficienty  $\beta_1, \dots, \beta_p$ , jejich standardní chyba SE a pomocí Waldovy statistiky je testována významnost těchto koeficientů v Coxově modelu. Je-li signifikance Waldova testu menší než 0,05, považujeme nezávisle proměnnou za efektivní prediktor sledovaného jevu. Veličina  $e^{\beta_i}$  ( $Exp(\beta_i)$ ),  $i = 1, \dots, p$  je hazard ratio (HR) a odhaduje procentuální změnu rizika sledovaného jevu (např. úmrtí) při jednotkové změně nezávisle proměnné.

Stejně jako u mnohonásobné lineární regrese a logistické regrese, kombinace regresních koeficientů  $\beta_1, \dots, \beta_p$  a hodnot odpovídajících faktorů je používána jako jediné číslo zvané prognostický index PI, který je dán vzorcem

$$PI = \beta_1 X_1 + \dots + \beta_p X_p. \quad (4.2)$$

Hodnota HR i hodnota funkce přežití (pravděpodobnost přežití) závisí pouze na celkové hodnotě PI (a samozřejmě i na čase), nikoli na hodnotách jednotlivých faktorů, což dokládá pro funkci hazardu následující rovnice

$$h(t) = h_0(t) e^{(\beta_1 X_1 + \dots + \beta_p X_p)} = h_0(t) e^{PI} \quad (4.3)$$

Pro funkci přežití je možné z předchozí rovnice a ze známého vztahu mezi funkcí přežití a funkcí hazardu odvodit

$$S(t) = e^{-\int_0^t h(u) du} = e^{-\int_0^t h_0(u) e^{PI} du} = e^{\left[ e^{PI} \left( -\int_0^t h_0(u) du \right) \right]} = \left( e^{-\int_0^t h_0(u) du} \right)^{e^{PI}} = S_0(t) e^{PI}, \quad (4.4)$$

tedy

$$S(t) = S_0(t) e^{PI}. \quad (4.5)$$

Z uvedeného vyplývá, že z Coxova regresního modelu pro funkci hazardu je možno přejít k regresnímu modelu pro funkci přežití s využitím stejných koeficientů pro uvažované faktory. Tímto postupem (který je obvykle součástí statistických programů) lze získat průběhy funkce přežití upravené pro určité hodnoty (kombinaci hodnot) sledovaných faktorů.



## 4.2 ML (maximum likelihood) odhady regresních koeficientů Coxova modelu

### 4.2.1 Likelihood funkce

Tento stručný výklad je zde uveden z důvodu, že likelihood funkce má klíčovou úlohu při teoretickém odvozování ML odhadů parametrů Coxova modelu.

Předpokládejme, že známe pravděpodobnostní rozložení sledované náhodné veličiny, které závisí na parametru (parametrech)  $\theta$ . Jedná se podmíněné rozložení  $f(x|\theta)$ , za podmínky že známe parametr  $\theta$ . Jedno pozorování jevu má hustotu  $f(x_i|\theta)$ .

Pro nezávislá pozorování je společné rozložení násobkem jednotlivých hustot nebo pravděpodobnostních funkcí:

$$f(x_1, x_2, \dots, x_n | \theta) = \prod_{i=1}^n f(x_i | \theta) \quad (4.6)$$

Můžeme interpretovat  $f(x_1, x_2, \dots, x_n | \theta)$  jako pravděpodobnost dané konfigurace pozorování, jestliže známe parametr  $\theta$ .

Věrohodnostní funkce je úměrná společnému rozložení:

$$L(x_1, x_2, \dots, x_n | \theta) \propto f(x_1, x_2, \dots, x_n | \theta) \quad (4.7)$$

Když mluvíme o rozložení, považujeme parametr za fixní a pozorování se mění. Jestliže mluvíme o věrohodnosti, pak jsou pozorování fixní a parametr se může měnit :

$$L(\theta | x_1, x_2, \dots, x_n) \quad (4.8)$$

Princip maximální věrohodnosti říká, že máme zvolit jako odhad takovou hodnotu parametru, která maximalizuje věrohodnost toho, že napozorujeme danou konfiguraci pozorování:

$$L(x_1, x_2, \dots, x_n | \hat{\theta}) \geq L(x_1, x_2, \dots, x_n | \theta) \quad (4.9)$$

Jestliže maximalizujeme funkci, která má derivaci, pak lze derivaci položit rovnou nule:

$$\frac{dL(x_1, x_2, \dots, x_n | \theta)}{d\theta} = 0 \quad (4.10)$$

Řešení této rovnice obdržíme hodnoty proměnné  $\theta$ .

Místo věrohodnostní funkce často derivujeme logaritmus věrohodnostní funkce, protože ten je rostoucí funkcí a maxima obou funkcí jsou stejná:

$$l(x_1, x_2, \dots, x_n | \theta) = \ln(L(x_1, x_2, \dots, x_n | \theta)) = \sum_{i=1}^n \ln(f(x_i | \theta)) \quad (4.11)$$

Práce se součtem je výhodnější než s násobky.

#### 4.2.2 ML odhady

Teorie ML odhadů je využita při hledání koeficientů Coxova modelu. Tyto koeficienty (parametry) se v obecném Coxově modelu značí  $\beta_i$ , jejich odhady se nazývají ML (maximum likelihood) odhady a značíme je  $\hat{\beta}_i, i = 1, \dots, n$ .

Stejně jako v logistické regresi, také ML odhady parametrů v Coxově odhadu jsou získány pomocí maximalizace tzv. likelihood funkce, obvykle značené  $L$ . Likelihood funkce je pravděpodobnost, že získáme naměřená data přežití za předpokladu, že platí obecný Coxův regresní model s neznámými parametry  $\beta_i, i = 1, \dots, n$ .

Jinými slovy, likelihood funkce udává pravděpodobnost, že při opakování studie přežití se skupinou jedinců náhodně vybraných z populace, získáme totožné výsledky pro funkci hazardu.

Likelihood funkce pro Coxův regresní model se stanoví obdobným postupem, který byl uveden v předchozí kapitole. V tomto konkrétním případě odpovídá pravděpodobnostnímu rozložení funkce hazardu a parametr  $\theta$  je nahrazen regresními koeficienty  $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_n$ . Odvození likelihood funkce pro Coxův regresní model a její přesný tvar je uveden v literatuře [9]. Tato funkce je součástí statistických počítačových programů.

Výraz pro likelihood funkci v Coxově regresním modelu je nazýván částečnou likelihood funkcí. Důvodem je fakt, že výraz bere do úvahy pravděpodobnosti selhání jen těch pozorovaných subjektů, u kterých bylo pozorováno selhání.

Tato částečná likelihood funkce může být zapsána jako součin dílčích likelihood funkcí, z nichž každá odpovídá jenomu selhání.

Výsledná likelihood funkce má následující tvar

$$L = L_1 x L_2 x L_3 x \dots x L_k = \prod_{j=1}^k L_j, \quad (4.12)$$

kde  $L_j$  odpovídá věrohodnosti selhání v čase  $t_j$ .

Je zaveden pojem skupina v riziku  $R(t_{(j)})$ , který označuje skupinu pozorovaných jedinců vystavených riziku selhání (tj. ještě žijících) v okamžiku  $j$ -tého pozorovaného selhání, tato skupina se stále zmenšuje.

Zavedení pojmu částečné likelihood funkce neznamená, že na její tvar cenzorování jedinci nemají žádný vliv. Tito jedinci sice nepřispívají k výslednému tvaru žádnou dílčí funkcí  $L_j$ , ale jejich existence v časech selhání před jejich cenzorováním ovlivňuje skupinu v riziku  $R(t_{(j)})$  a také hodnotu funkce hazardu v daném čase  $t_{(j)}$ .

Jakmile známe tvar likelihood funkce pro daný Coxův regresní model, můžeme najít hodnoty ML odhadů parametrů  $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_n$ , pro které daná likelihood funkce nabývá maxima. Obvykle je snadnější hledat maximum logaritmu likelihood funkce než funkce samé. Maximum likelihood funkce je největším z lokálních maxim, která je možno nalézt pomocí řešení soustavy rovnic, která vznikla parciální derivací likelihood funkce (resp. jejího logaritmu) postupně podle všech neznámých parametrů Coxova regresního modelu:

$$\frac{\partial L}{\partial \beta_i} = 0, \quad i = 1, \dots, p, \quad \text{kde } p \text{ je počet parametrů (koeficientů)}. \quad (4.13)$$

Řešení takto získané soustavy rovnic pro neznámé  $\beta_i, i = 1, \dots, n$  se obvykle provádí numerickými metodami. K řešení se dospěje v  $n$  iteračních krocích, kdy na vstupu prvního kroku jsou odhady hledaných parametrů a vstupy všech dalších kroků tvoří výstupy z kroku předchozího.

### 4.3 Výběr Coxova modelu

Numerické metody, jejichž princip byl přiblížen v předchozí kapitole umožňují získat vlastní odhad koeficientů, standardní chyby těchto odhadů a HR odpovídající jednotkové změně hodnoty daného faktoru. Waldova statistika je definována jako podíl odhadu koeficientu a jeho standardní chyby a má standardní normální rozložení  $Z$ , ze kterého je možno stanovit příslušnou dosaženou hladinu signifikance  $p$ .

$LR$  je vypočítáno podle vztahu  $LR = -2 \log L$ , kde  $L$  je hodnota likelihood funkce pro odhadnuté regresní koeficienty.

Výstupem z Coxovy regrese je obecně  $n$  matematických modelů, které zahrnují všechny možné kombinace zvolených faktorů i jejich interakcí definovaných součinem původních

faktorů. Všechny tyto modely respektují vzorec obecného Coxova regresního modelu, liší se jen kombinací uvažovaných faktorů.

Dalším krokem v celkovém algoritmu Coxovy regrese je posouzení obdržných modelů a výběr toho, který obsahuje jen statisticky významné faktory, popř. i jejich statisticky významné interakce (definované příslušným součinem primárních faktorů). Výběr modelu je prováděn iterativně tak, že se porovnává model se svým nejbližše předcházejícím modelem (tím, který obsahuje právě o jeden základní či interakční faktor méně) a na základě několika kritérií se rozhodne, zda i předcházející (o jeden faktor jednodušší) model vyhovuje, tedy zda je možno rozdílový faktor z modelu vypustit či ne. Iterace začíná v nejobecnějším modelu Coxovy regrese (se všemi faktory i jejich interakcemi).

V případě interakčních faktorů existují dvě nezávislé metody, jak rozhodnout o statistické významnosti a tedy i ponechání tohoto faktoru v modelu. První metoda se opírá o hodnotu signifikance  $p$  Waldovy statistiky interakčního faktoru. Hodnoty vyšší než 0,05 indikují nevýznamnost posuzovaného interakčního faktoru. Druhá metoda spočívá v nalezení rozdílu hodnot  $LR$  pro nejbližší předchozí model (bez interakčního faktoru) a uvažovaný model (s interakčním faktorem) a odečtením  $p$  hodnoty z chí-kvadrát statistiky. Hodnota vyšší než 0,05 opět indikuje nevýznamnost posuzovaného interakčního faktoru, který může být z modelu vypuštěn. Obě metody by měly poskytovat velmi podobné, ačkoli nikoli totožné výsledky.  $LR$  statistika má však lepší statistické vlastnosti a bývá proto upřednostňována.

O statistické významnosti přímých faktorů rozhoduje opět především hodnota signifikance  $p$  Waldovy statistiky, kde hodnoty vyšší jak 0,05 indikují statistickou nevýznamnost. Další metoda eliminace spočívá v porovnání HR hodnot v nejbližším předchozím modelu (bez uvažovaného přímého faktoru) a v uvažovaném modelu (s uvažovaným přímým faktorem). Pokud se HR hodnoty odpovídajících přímých faktorů v obou modelech zřetelně liší, uvažovaný přímý faktor by měl být v modelu ponechán. Dokonce i v případě dobré shody HR hodnot v obou modelech může existovat důvod pro ponechání uvažovaného přímého faktoru, a to tehdy, jsou-li intervaly spolehlivosti HR hodnot zřetelně užší v modelu s ponechaným přímým faktorem.

## 4.4 Výpočet hazard ratio ( $HR$ )

Obecně je  $HR$  definováno jako hazard (funkce hazardu) odpovídající určitému jednotlivci dělený hazardem (funkcí hazardu) příslušnou jinému jednotlivci. Předpokládá se, že oba jednotlivci sdílejí stejnou funkci hazardu resp. její Coxův regresní model, ale že se liší v hodnotách regresních faktorů.

Pro odhad  $HR$  lze tedy definovat vztah

$$\hat{HR} = e^{\hat{\beta}} \quad (4.14)$$

$$\hat{HR} = \frac{\hat{h}(t, X^*)}{\hat{h}(t, X)}, \quad (4.15)$$

kde  $X^* = (X_1^*, X_2^*, \dots, X_p^*)$  a  $X = (X_1, X_2, \dots, X_p)$

Dosazením Coxova regresního modelu do tohoto vztahu lze získat vzorec pro odhad  $HR$ , který závisí pouze na regresních koeficientech a hodnotách faktorů pro oba porovnávané jednotlivce, neboť funkce základního hazardu  $h_0(t)$  je pro oba jednotlivce stejná a vykrátí se:

$$\hat{HR} = \frac{\hat{h}(t, X^*)}{\hat{h}(t, X)} = \frac{\hat{h}_0(t) e^{\sum_{i=1}^p \hat{\beta}_i X_i^*}}{\hat{h}_0(t) e^{\sum_{i=1}^p \hat{\beta}_i X_i}} = e^{\sum_{i=1}^p \hat{\beta}_i (X_i^* - X_i)} \quad (4.16)$$

Tento obecný vzorec je možno dále upravit a diskutovat pro několik typických případů.

První případ popisuje situaci, kdy máme pouze jediný faktor binárního typu, který nabývá hodnot 0 a 1. Často má tento faktor význam příslušnosti ke sledované skupině, např. 0 pro léčenou skupinu, 1 pro skupinu neléčenou, tedy

$X = (X_1, X_2, \dots, X_p) = (X_1)$ , kde  $X_1$  je binární veličina, nabývá hodnot (0, 1).

Položme  $X_1^* = 1, X_1 = 0$ , pak dosazením do obecného vztahu pro  $HR$  obdržíme

$$\hat{HR} = e^{\left[ \hat{\beta}_1 (X_1^* - X_1) \right]} = e^{\left[ \hat{\beta}_1 (1-0) \right]} = e^{\hat{\beta}_1}. \quad (4.17)$$

Druhý případ popisuje situaci s jedním faktorem binárního typu, který pro porovnávané jedince nabývá hodnot 0 a 1, a s libovolným počtem dalších faktorů libovolného typu, které nevykazují žádné vzájemné interakce. Hodnoty těchto dalších faktorů jsou sice neznámé, ale stejné pro oba jedince. Dosazením do obecného vztahu pro  $HR$  obdržíme totožný výraz jako v předchozím kroku, je však třeba si uvědomit, že regresní koeficient pro první faktor binárního typu se bude obecně lišit od hodnoty téhož koeficientu v prvním případě. Důvodem tohoto rozdílu je to, že koeficient binárního faktoru v druhém případě byl určen z odlišného Coxova regresního modelu, který oproti případu prvnímu bral do úvahy možný vliv dalších faktorů.

Třetí případ popisuje situaci s jedním faktorem binárního typu, který pro porovnávané jedince nabývá opět hodnot 0 a 1, a s libovolným počtem dalších faktorů libovolného typu, které vykazují vzájemné interakce s prvním binárním faktorem, nikoli však mezi sebou navzájem. Hodnoty těchto dalších faktorů jsou sice neznámé, ale stejné pro oba jedince. V tomto případě jsou hodnoty těchto dalších faktorů nutné pro stanovení  $HR$ , jak vyplývá z následujícího vzorce

$$\hat{HR} = e^{\left[ \hat{\beta}_+ + \sum \hat{\delta}_j W_j \right]}, \quad (4.18)$$

kde  $\hat{\beta}_+$  je koeficient.

## 4.5 Coxův model funkce přežití

Výše popsaná metoda Coxovy regrese poskytuje odhad funkce hazardu ve tvaru

$$h(t, X) = \hat{h}_0(t) e^{\sum \hat{\beta}_i X_i} \quad (4.19)$$

včetně známých hodnot regresních parametrů  $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_n$ .

Ze známého vztahu mezi funkcí přežití  $S(t)$  a funkcí hazardu  $h(t)$  můžeme dosazením Coxova modelu pro funkci hazardu získat odhad Coxova modelu i pro funkci přežití ve tvaru

$$\hat{S}(t, X) = [S_0(t)] e^{-\sum \hat{\beta}_i X_i}. \quad (4.20)$$

K získání kompletního Coxova modelu pro funkci přežití nyní zbývá určit odhad základní funkce přežití  $\hat{S}_0(t)$ . Tento výpočet je proveden na základě úvahy, že z naší konkrétní studie známe hodnoty faktorů i jim odpovídající hodnotu  $S(t)$  určenou Kaplan-Meierovou křivkou. Dosadíme-li do známého Coxova modelu (se známými koeficienty faktorů) příslušné hodnoty faktorů, pak pro známé  $S(t)$  můžeme určit hledané  $S_0(t)$ . Určení  $S_0(t)$  založené na tomto postupu je součástí používaných statistických softwarových programů. Stanovením  $S_0(t)$  se odhad Coxova modelu funkce přežití stává kompletním a můžeme jej použít ke stanovení funkce přežití pro libovolnou kombinaci hodnot faktorů.

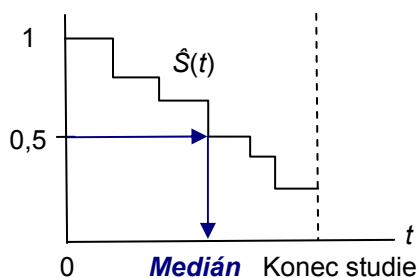
## 5 Medián a průměr v analýze přežívání

Výpočet popisných charakteristik mediánu a průměru pro data, která obsahují cenzorované údaje, se provádí specifickým způsobem, neboť nelze použít metody, které jsou známy ze zpracování pouze konečných necenzorovaných dat.

### 5.1 Medián

Medián přežití je definován jako čas na časové ose, kterému odpovídá pravděpodobnost přežití 0,5. Hodnota mediánu přežití se určí z tabulky přežití, která obsahuje pravděpodobnosti přežití v časech, ve kterých nastala událost, nebo se graficky určí z křivky přežití.

Obr. 5.1 Určení mediánu přežívání z křivky přežití.



### 5.2 Průměr

Průměrný čas přežití  $\mu$  je definován

$$\mu = EX = \int_0^{\infty} x dF(x) = \int_0^{\infty} S(x) dx \quad (5.1)$$

Pro výpočet odhadu průměrného času přežití se používá několika metod. Při numerickém postupu, který využívá numerického integrování, vycházíme ze vztahu

$$\hat{\mu} = \sum_{i=1}^{k-1} \hat{S}(t_i^+) (t_{i+1} - t_i), \quad (5.2)$$

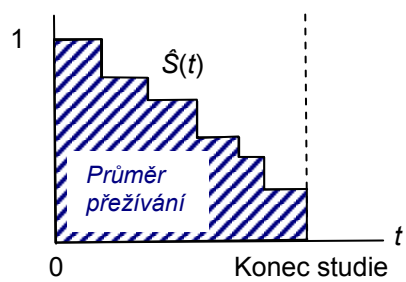
kde

$\hat{S}(t_i^+)$  je pravděpodobnost přežití v intervalu bezprostředně následujícím po čase  $t_i$ .



Průměrný čas přežívání je tedy plocha pod křivkou přežití.

Obr. 5.2 Grafický význam průměru přežívání.



## **6 Počítačová podpora**

### **6.1 Standardní aplikace**

Vstupní data pro analýzu přežívání obvykle pořizujeme v tabulkovém procesoru MS Excel. Data jsou obvykle zadávána uživateli z řad lékařů Lékařské fakulty a Fakultní nemocnice a MS Excel je program široce dostupný. Data jsou následně zpracována naším statistickým oddělením, který pro zpracování používá program SPSS verze 15. Pro potřeby analýzy přežívání tento program poskytuje následující možnosti:

- odhaduje průběh funkce přežívání Kaplan-Meierovou metodou včetně grafického znázornění
- výpočítává popisné charakteristiky - průměr a medián přežívání
- porovnává více křivek přežití třemi statistickými testy (log-rank test, Tarone-Ware test a Wilcoxonův test)
- provádí Coxovu regresní analýzu.

### **6.2 Vlastní programové vybavení**

Protože standardní statistický program SPSS neumožňuje stanovit odhad křivky přežívání včetně intervalů spolehlivosti, což bývá častý požadavek lékařských uživatelů, byla navržena a implementována aplikace, která z dat získaných z SPSS tuto funkcionalitu poskytuje.

Vstupem této aplikace je popis jedné nebo více křivek přežívání ve formátu XML. Tento formát dat získáme tak, že exportujeme z programu SPSS tabulku přežívání do formátu XLS a následně tato data programem Excel uložíme ve formátu XML.

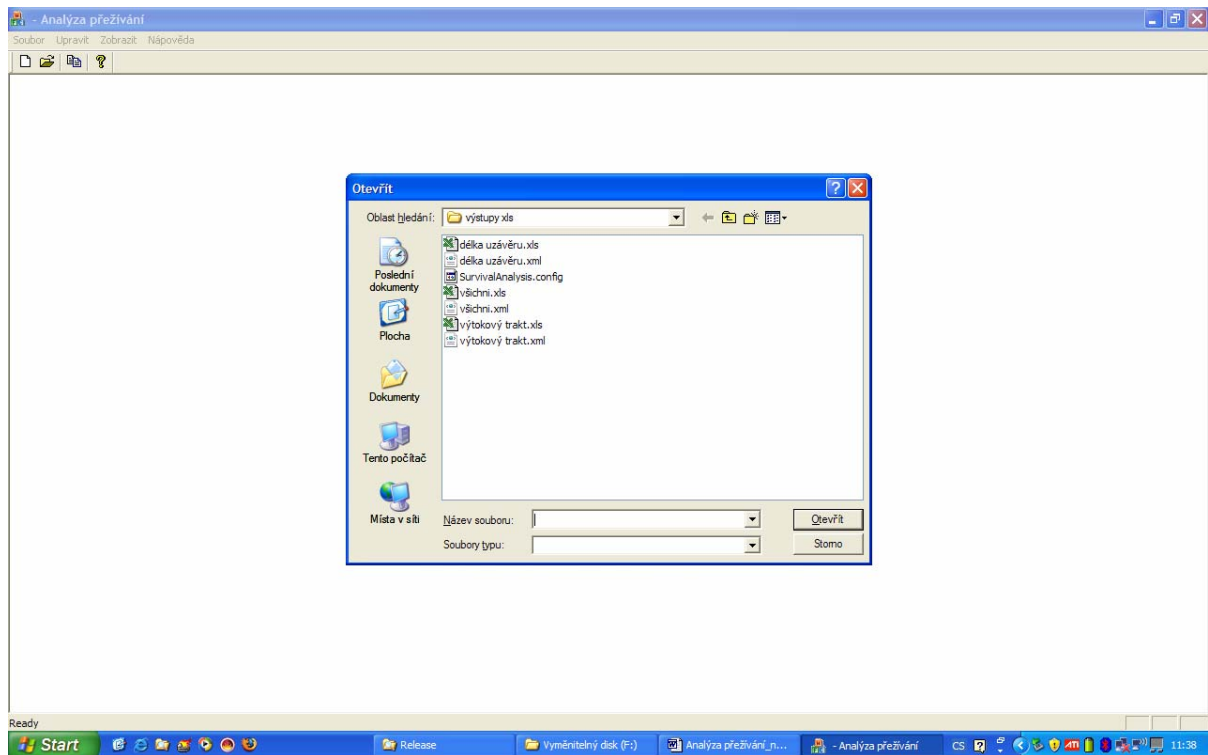
Aplikace poskytuje dvě základní funkce, které chybí v programu SPSS. První funkce umožňuje zobrazit jednu nebo více křivek přežití ze vstupního XML souboru včetně 95% intervalu spolehlivosti. Druhá funkce spočívá v možnosti určit z uvedených křivek přežití pravděpodobnost přežití včetně intervalů spolehlivosti v libovolně zvolených časech a tato data zobrazit jako tabulku.

Obě tyto funkce je možno v určitém rozsahu konfigurovat pomocí dialogů nastavení v aplikaci. Například u první funkce je možno interaktivně volit vzhled křivek přežití (barva, tloušťka, typ čáry, typ značek pro cenzorované údaje atd.), popisky a měřítko na osách grafu.

U druhé funkce je samozřejmě nutno zadat, pro které časy je požadován odečet pravděpodobnosti přežití včetně intervalů spolehlivosti, a dále můžeme konfigurovat vzhledové vlastnosti tabulek.

Soubor ve formátu XML s daty je možno otevřít příkazem „Otevřít“ z menu nebo z panelu nástrojů viz obr. 6.1.

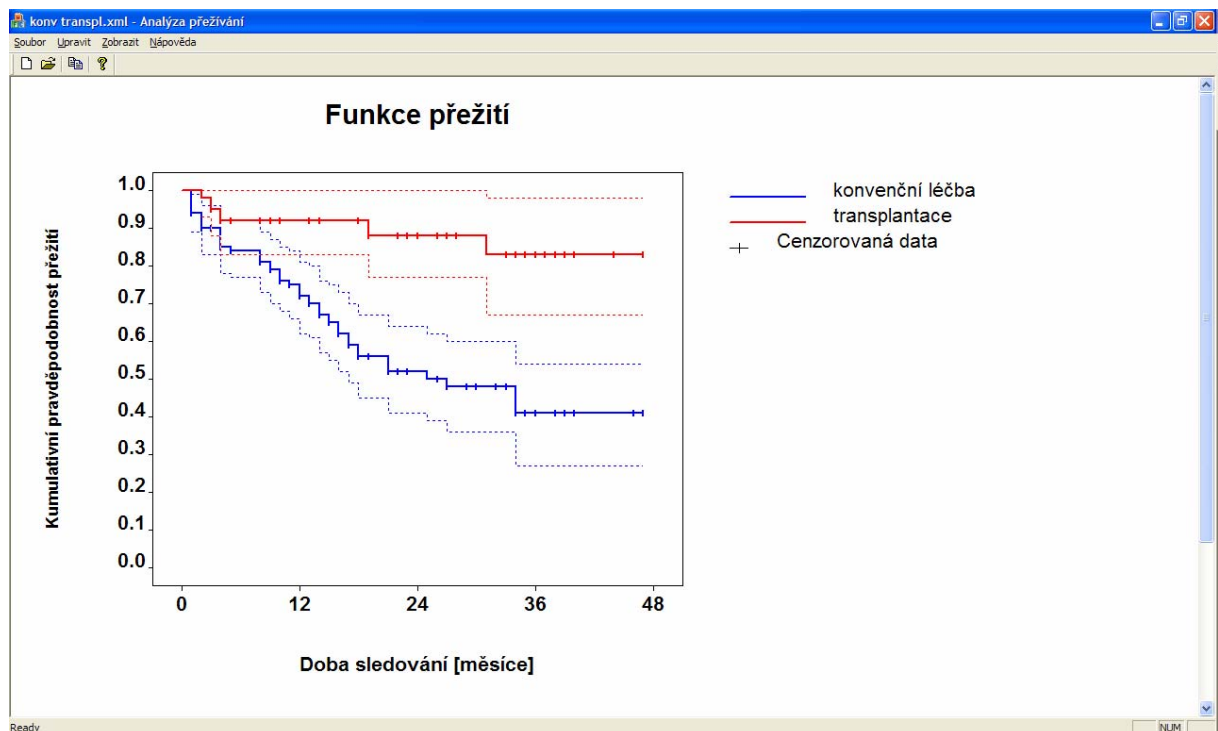
Obr. 6.1



Poslední čtyři otevřené soubory je možno rovněž rychle otvírat standardním způsobem z menu „Soubor“.

Po otevření souboru se vykreslí graf křivek přežití způsobem, který je vidět na obr. 6.2.

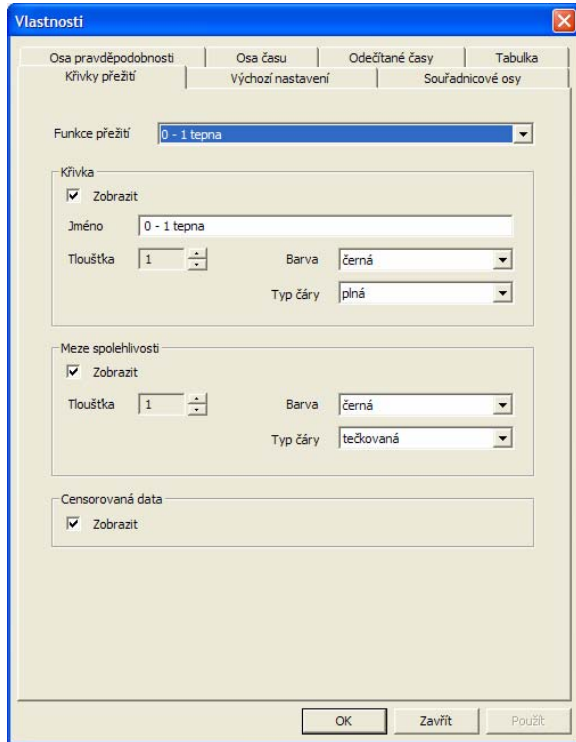
Obr. 6.2



Vlastnosti zobrazovaných křivek lze interaktivně zadávat v dialogovém okně, které se otevře příkazem „Vlastnosti“ z menu „Soubor“.

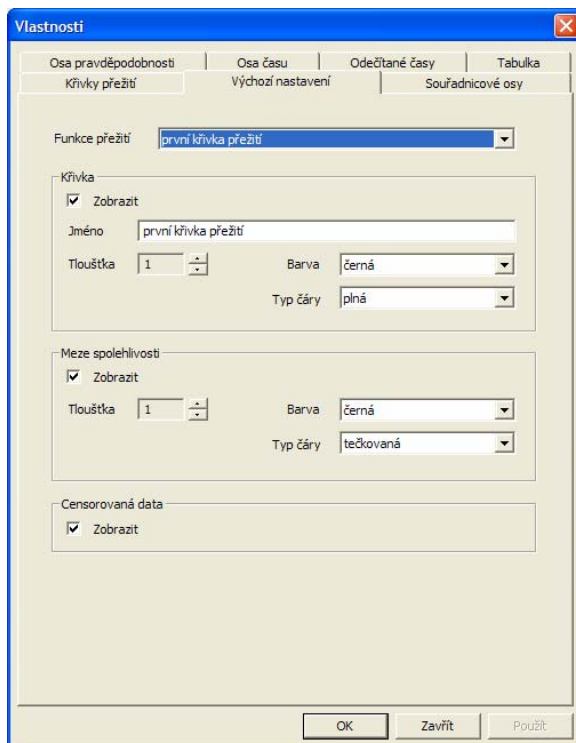
Karta „Křivky přežití“ zadává vlastnosti křivek a jejich intervalu spolehlivosti viz obr. 6.3.

Obr. 6.3



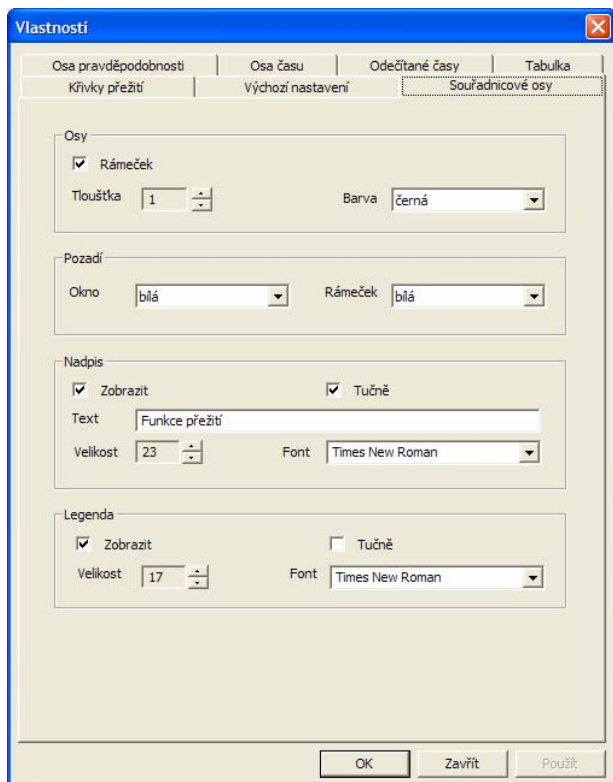
Karta „Výchozí nastavení“ zadává počáteční vlastnosti křivek v nově otevřeném souboru viz. obr. 6.4.

Obr. 6.4



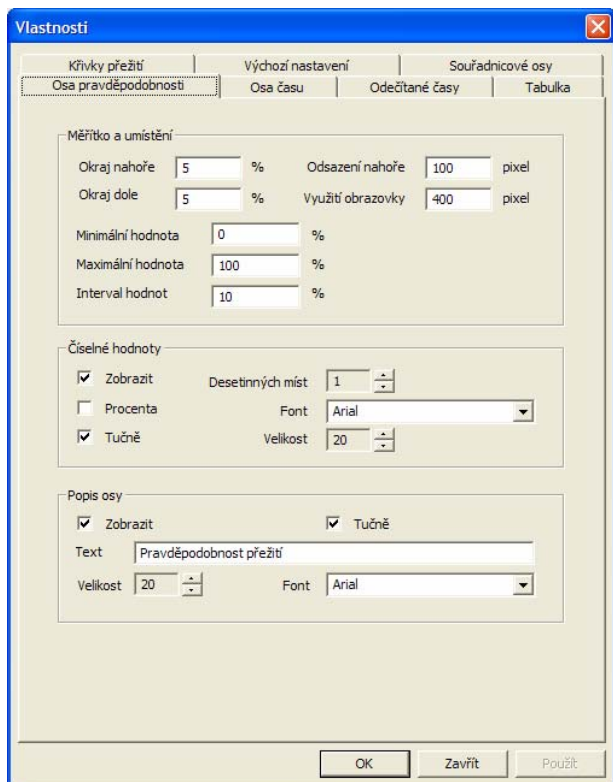
Karta „Souřadnicové osy“ umožňuje zadávat vlastnosti souřadnicových os viz obr. 6.5.

Obr. 6.5



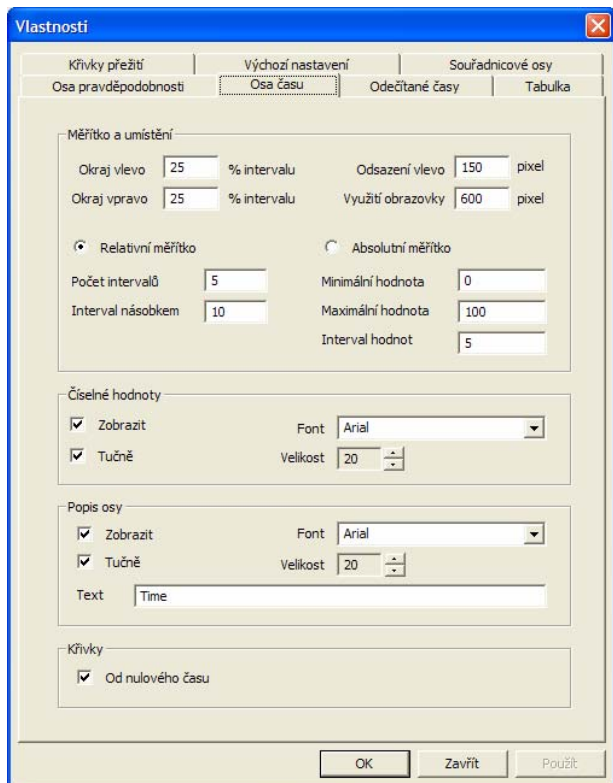
Karta „Osa pravděpodobnosti“ slouží k nastavení svislé osy, zejména měřítka a zobrazovaného rozsahu viz obr. 6.6.

Obr. 6.6



Karta „Osa času“ obdobně slouží k nastavení vodorovné osy, opět zejména měřítka a zobrazovaného rozsahu viz obr. 6.7.

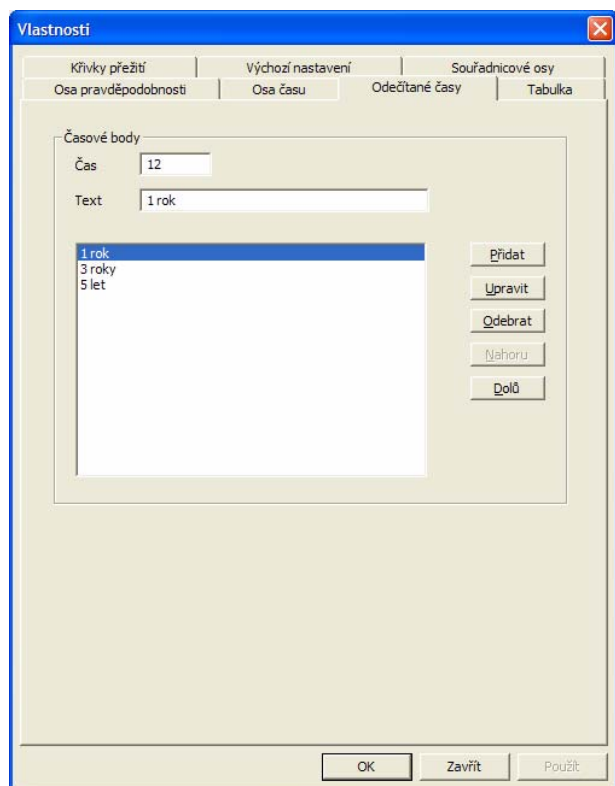
Obr. 6.7



Druhou hlavní funkcí aplikace je tabulkové zobrazení pravděpodobnosti přežití včetně 95% intervalu spolehlivosti ve zvolených časech. Vyvolání této funkce včetně její konfigurace je možno provést opět z dialogového okna „Vlastnosti“, které se otvírá příkazem vlastnosti z menu „Soubor“.

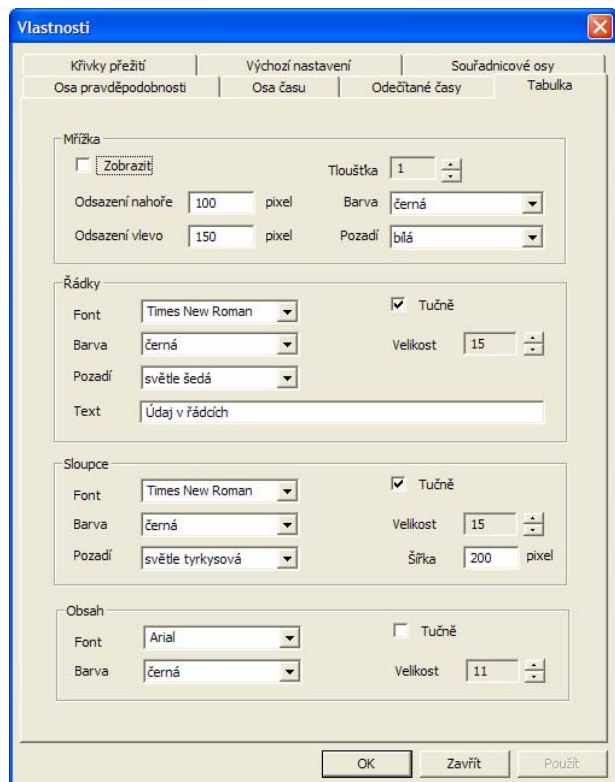
Na kartě „Odečítané časy“ je možno zadat libovolný počet časových bodů, pro které bude tabulka sestavena v dolní části pracovního prostoru viz obr.6.8.

Obr. 6.8



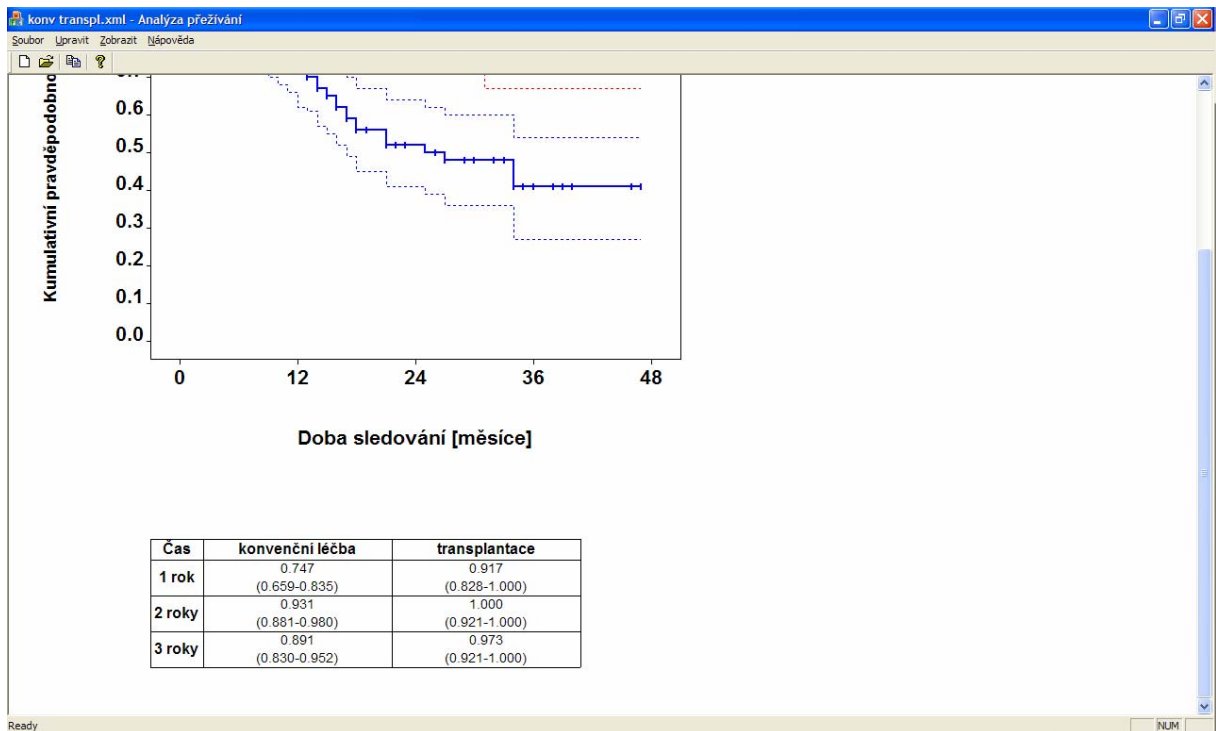
Na kartě „Tabulka“ se zadávají zobrazovací vlastnosti tabulky odečítaných časů viz obr. 6.9.

Obr. 6.9

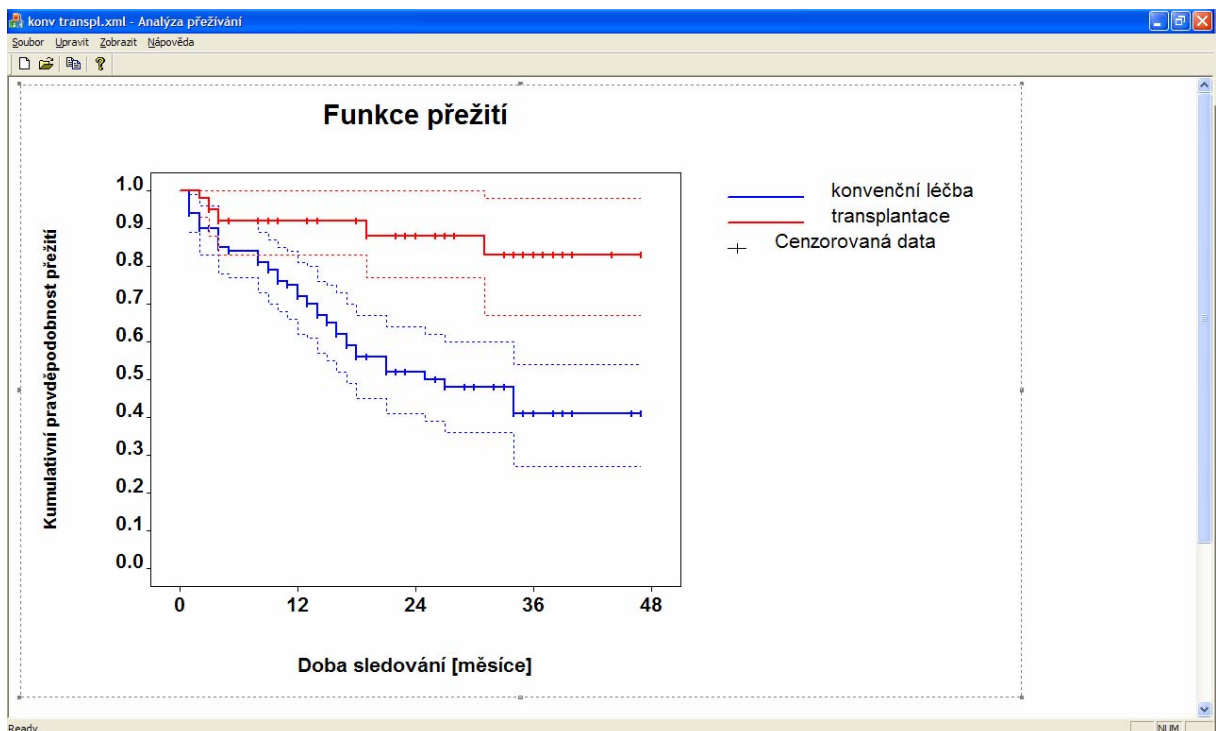


Po zadání časových bodů a po nastavení zobrazovacích vlastností tabulky se požadovaná tabulka zobrazí na obrazovce pod grafem viz obr. 6.10.

Obr. 6.10



Libovolnou část pracovního prostoru, který může obsahovat jak graf křivek, tak i tabulku odečítaných časů lze pomocí myši označit a zkopírovat do schránky příkazem „Kopírovat“ z menu „Úpravy“. Na obr. 6.11 je znázorněno označení grafu, který je možno zkopírovat do schránky a následně vložit do jiného souboru např. do dokumentu v programu MS Word.





## **7 Příklady**

### **7.1 Studie přežívání hemato-onkologických pacientů**

V klinické studii, která byla provedena ve spolupráci s Hemato-onkologickou klinikou LF a FN v Olomouci, bylo sledováno přežívání hemato-onkologicky nemocných pacientů. Skupinu tvořilo 176 pacientů s hematologickými diagnózami (přes 95 % pacientů mělo hematologické nádory), kteří podstoupili transplantaci krvetvorných kmenových buněk (TKB). U pacientů bylo sledováno mnoho dalších faktorů (viz tabulky 7.1a, 7.1b). Cílem studie bylo zjistit, které faktory jsou statisticky významné pro přežívání a určit míru, s jakou jednotlivé faktory přežívání ovlivňují.

Tab. 7.1. a Vstupní data

číslo pacienta	příjemce dárce		věk příjemce dárce v letech	věk dárce v letech	diagnóza akutní leukémie	příbuzný dárce	nemýelo ablativní režim	ATG nebo Campath 1H v přípravném režimu	GvHD- profylaxe- počet IS	CD34+	rejekce štěpu po		akutní GvHD	chronická GvHD	úmrtí po TKB	dobu sledování v měs.
	muž	muž									TKB	TKB				
1	1	1	21	24	1	1	0	0	2	5,30	0	0	0	0	1	10,03
2	1	0	41	53	0	1	1	0	1	7,36	0	0	0	0	0	87,77
3	1	1	44	46	1	1	1	1	1	4,55	0	0	0	0	1	4,26
4	1	1	24	20	1	1	0	0	2	6,70	0	0	0	0	1	10,20
5	1	0	52	50	1	1	0	0	2	6,52	0	1	0	0	1	3,57
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
170	1	0	39	24	0	0	0	1	2	15,6	0	1	0	0	1	21,61
171	1	0	60	29	1	0	1	0	2	6,9	0	1	0	0	0	29,15
172	1	1	54	35	1	0	0	1	2	7,5	0	0	0	0	0	29,34
173	1	0	54	36	1	0	0	0	2	10,7	0	1	1	1	1	18,03
174	1	1	52	37	0	0	0	1	3	5,9	0	1	0	0	1	3,34
175	0	1	42	37	0	0	0	1	3	3,6	0	0	0	0	1	0,69
176	0	1	48	20	1	0	0	0	3	5,6	0	0	0	0	1	14,00

Vysvětlivky

příjemce muž 1 – ano, 0 – ne  
dárce muž 1 – ano, 0 – ne  
diagnóza akutní leukémie 1 – ano, 0 – ne  
příbuzný dárce (většinou sourozenec) 1 – ano, 0 – ne  
nemýeloablativní režim (přípravný režim před transplantací) 1 – ano, 0 – ne  
ATG (antithymocytární globulin) nebo Campath-1H (specifické protilátky použité v přípravném režimu) 1 – ano, 0 – ne  
GvHD (nemoc štěpu proti hostiteli Graft-Versus-Host Disease) profylaxe - vyjádřeno počtem IS (imunosupresiv)  
CD34+ počet kmenových buněk v 10<sup>6</sup> na 1 kg  
rejekce štěpu po TKB (transplantace kostních buněk) 0 – přihojí se, 1 – nepřihojil  
akutní GvHD 1 – ano, 0 – ne  
chronická GvHD 1 – ano, 0 – ne  
úmrtí po TKB 1 – ano, 0 – ne

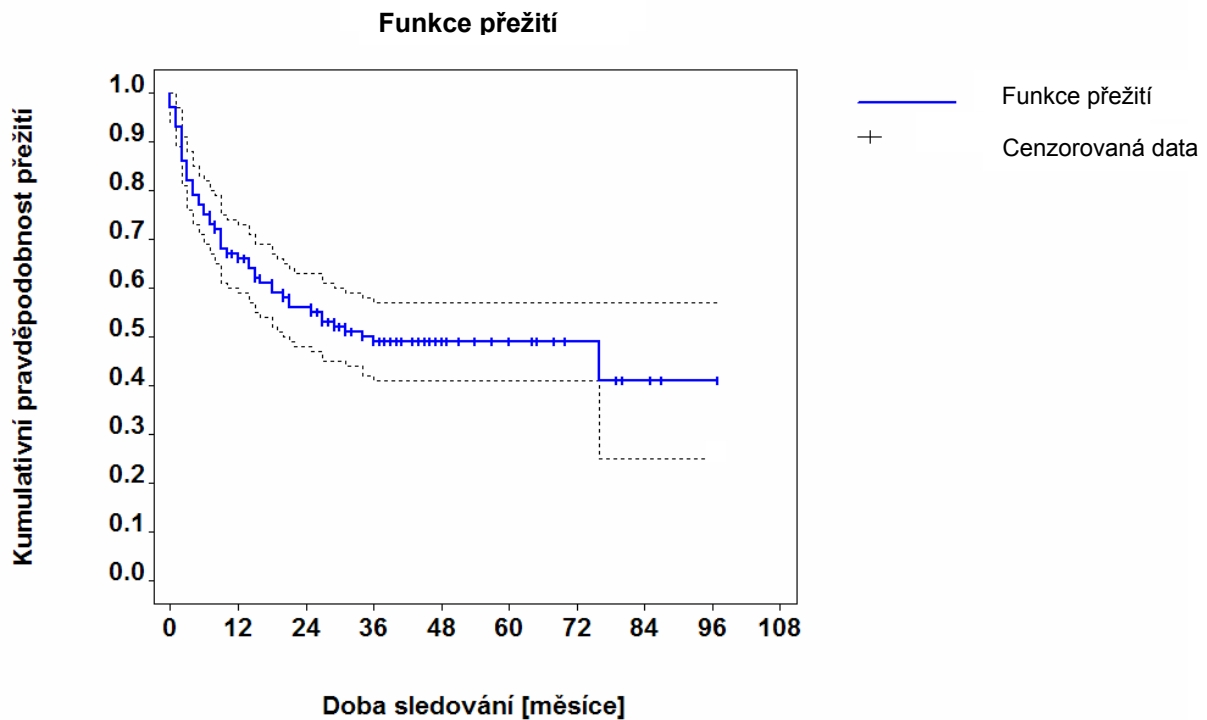
Tab. 7.1. a Vstupní data - pokračování

číslo pacienta	přítomnost IL-6-174*G alely u pacienta	přítomnost IL-6-174*G alely u pacienta	přítomnost MCP-1-2076*T alely u pacienta	přítomnost MCP-1-2076*T alely u pacienta	přítomnost IL-6-174*G alely u dárce	přítomnost IL-6-174*G alely u dárce	přítomnost MCP-1-2518*G alely u dárce	přítomnost MCP-1-2518*G alely u dárce	počet MCP-1-2076*T alely u dárce
1	0	0	1	0	1	1	1	0	0
2	0	0	0	1	1	1	1	0	0
3	1	1	0	1	1	1	1	0	1
4	0	0	0	0	0	0	0	1	0
5	1	1	0	0	1	1	1	0	0
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
170	0	0	1	0	1	1	1	1	0
171	1	1	0	0	0	0	1	0	0
172	1	1	0	1	1	1	1	1	1
173	1	1	1	1	1	1	1	0	0
174	1	1	1	1	1	1	1	1	0
175	0	0	0	0	1	1	1	0	0
176	0	0	1	1	1	1	1	1	1

Vysvětlivky

přítomnost určitého typu alely u pacientů a dárců 1 – přítomna, 0 - nepřítomna

Kaplan-Meierova funkce přežití včetně 95% CI má tvar



Odhad pravděpodobnosti přežití v zadaných časech:

Čas	Pravděpodobnost přežití (95% CI)
1 rok	0.661 (0.591-0.732)
3 roky	0.499 (0.418-0.579)
5 let	0.487 (0.406-0.569)
7 let	0.406 (0.246-0.567)

**Means and Medians for Survival Time**

Mean <sup>a</sup>				Median			
Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
		Lower Bound	Upper Bound			Lower Bound	Upper Bound
51,270	3,754	43,912	58,628	34,098	15,132	4,440	63,756

a. Estimation is limited to the largest survival time if it is censored.

Odhad průměrného přežívání je 51,3 měsíců (95% CI 43,9 – 58,6 měs.), odhad mediánu přežívání je 34,1 měsíců (95% CI 4,4 – 63,8 měs.)

Coxova regrese – počítačové zpracování

**Case Processing Summary**

		N	Percent
Cases available in analysis	Event <sup>a</sup>	83	47,2%
	Censored	93	52,8%
	Total	176	100,0%
Cases dropped	Cases with missing values	0	,0%
	Cases with negative time	0	,0%
	Censored cases before the earliest event in a stratum	0	,0%
	Total	0	,0%
Total		176	100,0%

a. Dependent Variable: OS - time [months]

Tabulka obsahuje přehled 176 případů (pacientů), z nichž 93 dat bylo cenzorováno a 83 pacientů zemřelo. Žádný pacient nebyl ze zpracování vyřazen.

Variables not in the Equation<sup>a</sup>

	Score	df	Sig.
přijemcemuž	,242	1	,623
dárcemuž	,042	1	,838
věkpřjemce	1,016	1	,313
věkdárce	2,206	1	,137
diagnózaakutníleukémie	3,183	1	,074
nemyeloablativnírežim	,035	1	,851
ATGneboCampath1 Hvpřípravnémrežimu	13,398	1	<b>,0003</b>
GvHDprofylaxepočetIS	1,181	1	,277
CD34x106kg	,377	1	,539
rejekceštěpupoTKB	4,070	1	<b>,044</b>
akutníGvHD	14,500	1	<b>,0001</b>
chronickáGvHD	18,524	1	<b>&lt;0,0001</b>
přítomnostIL6174 Galelyupacienta	,003	1	,954
přítomnostIL6nt565 Galelyupacienta	,111	1	,740
přítomnostMCP12518 Galelyupacienta	4,553	1	<b>,033</b>
přítomnostMCP12076 Talelyupacienta	,715	1	,398
přítomnostIL6174 Galelyudárce	1,868	1	,172
přítomnostIL6nt565 Galelyudárce	2,098	1	,147
přítomnostMCP12518 Galelyudárce	1,537	1	,215
přítomnostMCP12076 Talelyudárce	,410	1	,522

a. Residual Chi Square = 81,252 with 20 df Sig. = ,000

Tabulka uvádí všechny vstupní parametry s vypočítanými statistickými signifikancemi. Je zřejmé, že k významným prediktorům patří tyto veličiny:

- ATG nebo Campath 1 H v přípravném režimu
- rejekce štěpu po TKB
- akutní GvHD
- chronická GvHD
- přítomnost IL MCP 12076G alely u pacienta.

Variables in the Equation

		B	SE	Wald	df	Sig.	Exp(B)	95,0% CI for Exp(B)	
								Lower	Upper
Step 1	chronická GvHD	-1,166	,286	16,602	1	,000	,311	,178	,546
Step 2	akutní GvHD	1,016	,225	20,316	1	,000	2,763	1,776	4,298
	chronická GvHD	-1,346	,290	21,537	1	,000	,260	,148	,460
Step 3	ATG nebo Campath1H v přípravném režimu	-,999	,267	13,942	1	,000	,368	,218	,622
	akutní GvHD	,867	,228	14,516	1	,000	2,380	1,524	3,718
	chronická GvHD	-1,428	,289	24,343	1	,000	,240	,136	,423
Step 4	ATG nebo Campath1H v přípravném režimu	-,968	,269	13,000	1	,000	,380	,224	,643
	akutní GvHD	,952	,232	16,857	1	,000	2,592	1,645	4,084
	chronická GvHD	-1,521	,293	26,993	1	,000	,218	,123	,388
	přítomnost IL6174G alely u dárce	,732	,329	4,960	1	,026	2,079	1,092	3,961
Step 5	ATG nebo Campath1H v přípravném režimu	-,931	,269	11,981	1	,001	,394	,233	,668
	rejekceštěpupoTKB	1,170	,546	4,588	1	,032	3,222	1,104	9,399
	akutní GvHD	1,047	,241	18,931	1	,000	2,848	1,777	4,563
	chronická GvHD	-1,504	,295	26,061	1	,000	,222	,125	,396
	přítomnost IL6174G alely u dárce	,819	,333	6,033	1	,014	2,268	1,180	4,360
Step 6	ATG nebo Campath1H v přípravném režimu	-,986	,270	13,328	1	<b>,0003</b>	,373	,220	,633
	CD34x106kg	-,061	,030	4,145	1	<b>,0418</b>	,940	,887	,998
	rejekceštěpupoTKB	1,180	,549	4,622	1	<b>,0316</b>	3,253	1,110	9,536
	akutní GvHD	1,190	,247	23,284	1	<b>&lt;0,0001</b>	3,287	2,027	5,331
	chronická GvHD	-1,632	,300	29,504	1	<b>&lt;0,0001</b>	,196	,109	,352
	přítomnost IL6174G alely u dárce	,889	,337	6,959	1	<b>,0083</b>	2,433	1,257	4,709

Metodou postupného přidávání nezávisle proměnných do modelu bylo v šesti krocích postupně navrženo šest modelů. U proměnných zařazených do modelu byly odhadnuty koeficienty B, jejich chyby SE, hodnoty Waldovy statistiky (Wald), počet stupňů volnosti (df), signifikance (Sig.) a bylo spočítáno Exp(B), tedy hazard ratio HR (změna rizika exitu při jednotkové změně nezávisle proměnné) a 95% intervaly spolehlivosti pro HR.

### Závěr Coxovy regrese:

- Pokud byly v přípravném procesu použity specifické protilátky ATG (antithymocytární globulin) nebo Campath-1H, snižuje se riziko úmrtí po TKB 0,373x. (95% CI 0,22 – 0,633).
- Je-li počet kmenových buněk CD34 vyšší o  $10^6$  buněk na 1 kg, snižuje se riziko úmrtí po TKB 0,94x (95% CI 0,887 – 0,998).
- Pokud nedošlo k přihojení štěpu po TKB, zvyšuje se riziko úmrtí 3,25x (95% CI 1,11 – 9,54).
- Rozvoj akutní GvHD zvyšuje riziko úmrtí po TKB 3,29 x (95% CI 2,03 - 5,33).
- Rozvoj chronické GvHD snižuje riziko úmrtí po TKB 0,2x (95% CI 0,11 – 0,35).
- Přítomnost IL9174G alely zvyšuje riziko úmrtí po TKB 2,433x. (95% CI 1,26 – 4,71).

## **7.2 Studie přežívání pacientů s mnohočetným myelomem**

V klinické studii, která byla provedena ve spolupráci s III. interní klinikou LF a FN v Olomouci, bylo sledováno přežívání pacientů s mnohočetným myelomem. Skupinu tvořilo 138 pacientů, u pacientů byly sledovány další faktory (viz tabulka 7.2). Cílem studie bylo opět zjistit, které faktory jsou statisticky významné pro přežívání a určit míru, s jakou jednotlivé faktory přežívání ovlivňují.



Tab. 7.2 Vstupní data

Číslo pacienta	Skupina	Věk	Pohlaví	Stádium nemoci	Typ řetězce imunoglobulinů	Kreatinin	Beta2-mikroglobulin	ICTP	FAS	Řetězce kappa	Řetězce lambda	Poměr K/L	Úmrtí	Doba sledování [měsíce]
1	1	61	m	B	L	165	9,64	4,4	5160	9,2	272,7	0,03	1	16
2	1	69	m	A	K	97,0	2,98	4,4	8849	43,6	6,99	6,24	0	35
3	1	45	m	A	K	106	3,95	8,4	3223	58,76	6,61	8,90	1	11
4	1	75	m	A	K	77	4,97	14,2	8381	115,9	11,88	9,75	1	8
98	1	77	m	A	L	97	2,9	4,56	4048	12,94	369,5	0,04	0	3
99	1	63	ž	A	K	79	3,1	8,4					0	4
100	1	51	m	A	K	72	1	8,4					0	3
101	1	71	ž	A	L	64	2,9	14,2					0	2
103	2	57	ž	A	K	73,0	2,51	4	7571	45,12	7,16	6,30	0	44
104	2	49	ž	A	L	117	5,93	12,6	11008	6,85	1117	0,00	0	34
137	2	64	ž	A	L	80	3,28	7,73	46655	10,8	612,65	0,02	0	4
138	2	63	ž	A	K	70	3,55	16,3	10569	348,25	9,85	35,36	0	3

## Vysvětlivky

Skupina 1 - pacienti léčení konvenční chemoterapií, 2 - pacienti, kteří podstoupili autologní transplantaci kostní dřeně

Věk - věk pacientů v letech

Pohlaví m - muž, ž - žena

Stádium nemoci - stádium nemoci určené podle funkce ledvin A - normální funkce ledvin, B - porušená funkce ledvin

Typ řetězce imunoglobulinů K - kappa, L - lambda

Kreatinin - biochemický parametr měřený v umol/l

Beta2-mikroglobulin - prognostický faktor u nádorových onemocnění měřený v mg/l

ICTP - faktor remodelace kostní tkáně měřený v ug/l

FAS - povrchové molekuly zodpovědné za spouštění programované buněčné smrti měřené v ng/l

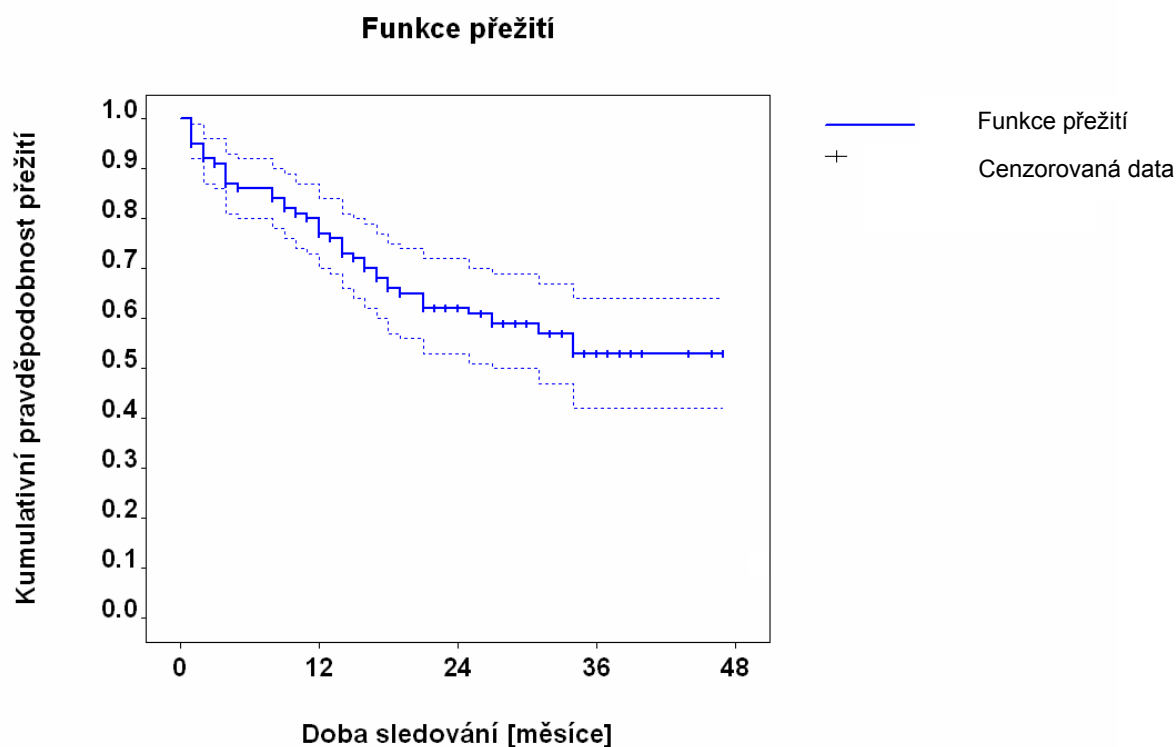
Řetězce kappa - kvantifikované lehké řetězce imunoglobulinů měřené v mg/l

Řetězce lambda - kvantifikované lehké řetězce imunoglobulinů měřené v mg/l

Poměr K/L - poměr kvantifikovaných řetězců kappa a lambda

Úmrtí 1 - došlo k úmrtí, 0 - čas sledování je cenzorován

Kaplan-Meierova funkce přežití včetně 95% CI pro celý soubor pacientů má tvar:



Odhad pravděpodobnosti přežití v zadaných časech:

Čas	Pravděpodobnost přežití (95% CI)
1 rok	0.792 (0.722-0.862)
3 roky	0.524 (0.414-0.634)

Coxova regrese – počítačové zpracování

**Case Processing Summary**

		N	Percent
Cases available in analysis	Even <sup>a</sup>	43	31,2%
	Censored	78	56,5%
	Total	121	87,7%
Cases dropped	Cases with missing values	17	12,3%
	Cases with negative time	0	,0%
	Censored cases before the earliest event in a stratum	0	,0%
	Total	17	12,3%
Total		138	100,0%

a. Dependent Variable: OS

Tabulka udává přehled 138 případů (pacientů), z nichž 78 dat bylo cenzorováno a 43 pacientů zemřelo. 17 případů bylo vypuštěno ze zpracování, neboť neměly kompletní údaje.

#### Variables not in the Equation<sup>a</sup>

	Score	df	Sig.
Skupina	10,617	1	<b>,0011</b>
Věk	6,105	1	<b>,0135</b>
Pohlaví	,001	1	,970
Stádium nemoci	12,424	1	<b>,0004</b>
Typ řetězce	2,454	2	,293
Kreatinin	10,942	1	<b>,0009</b>
beta2-mikroglobulin	20,047	1	<b>&lt;0,0001</b>
ICTP	17,707	1	<b>&lt;0,0001</b>
FAS	4,087	1	<b>,0432</b>
Řetězce kappa	,007	1	,934
Řetězce lambda	6,292	1	<b>,0121</b>
Poměr K/L	,047	1	,828

a. Residual Chi Square = 50,061 with 13 df Sig. = ,000

Tabulka uvádí všechny vstupní parametry s vypočítanými statistickými signifikancemi.

#### Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95,0% CI for Exp(B)	
							Lower	Upper
Step 1 beta2-mikro globulin	,050	,012	17,421	1	,000	1,051	1,027	1,076
Step 2 beta2-mikro globulin	,046	,012	14,394	1	,000	1,047	1,022	1,072
ICTP	,004	,002	6,620	1	,010	1,004	1,001	1,008
Step 3 skupina	-1,332	,537	6,138	1	<b>,013</b>	,264	,092	,757
beta2-mikro globulin	,037	,013	8,417	1	<b>,004</b>	1,037	1,012	1,063
ICTP	,004	,002	5,670	1	<b>,017</b>	1,004	1,001	1,007

#### Závěr Coxovy regrese:

Metodou postupného přidávání nezávisle proměnných do modelu bylo v šesti krocích postupně navrženo šest modelů. U proměnných zařazených do modelu byly odhadnuty koeficienty B, jejich chyby SE, hodnoty Waldovy statistiky (Wald), počet stupňů volnosti (df), signifikance (Sig.) a bylo spočítáno Exp(B), tedy hazard ratio HR (změna rizika exitu při jednotkové změně nezávisle proměnné) a 95% intervaly spolehlivosti pro HR.

Statisticky významné faktory pro predikci přežívání pacientů s mnohočetným myelomem zahrnují skutečnost, zda pacienti podstoupili autologní transplantaci kostní dřeně, hodnotu beta2-mikroglobulinu a faktor remodelace kostní tkáně ICTP.

Míra, s jakou tyto faktory ovlivňují přežívání, je následující:

- a) Pokud pacienti podstoupí autologní transplantaci kostní dřeně, snižuje se riziko úmrtí 0,264x. (95%CI 0,092 – 0,757).
- b) Zvýší-li se hodnota beta2-mikroglobulinu o 1 mg/l, zvýší se riziko úmrtí 1,037x (95%CI 1,012 – 1,063).
- c) Zvýší-li se hodnota ICTP o 1 ug/l, zvýší se riziko úmrtí 1,004x (95%CI 1,001 – 1,007).

Na základě těchto poznatků byla provedena Kaplan-Meierova analýza pro pacienty, kteří byli rozděleni podle

- a) příslušnosti ke skupině (pacienti léčení konvenčně a transplantovaní)
- b) hodnot beta2-mikroglobulinu (pacienti s nižší a vyšší hladinou)
- c) ICTP (pacienti s nižší a vyšší hladinou)

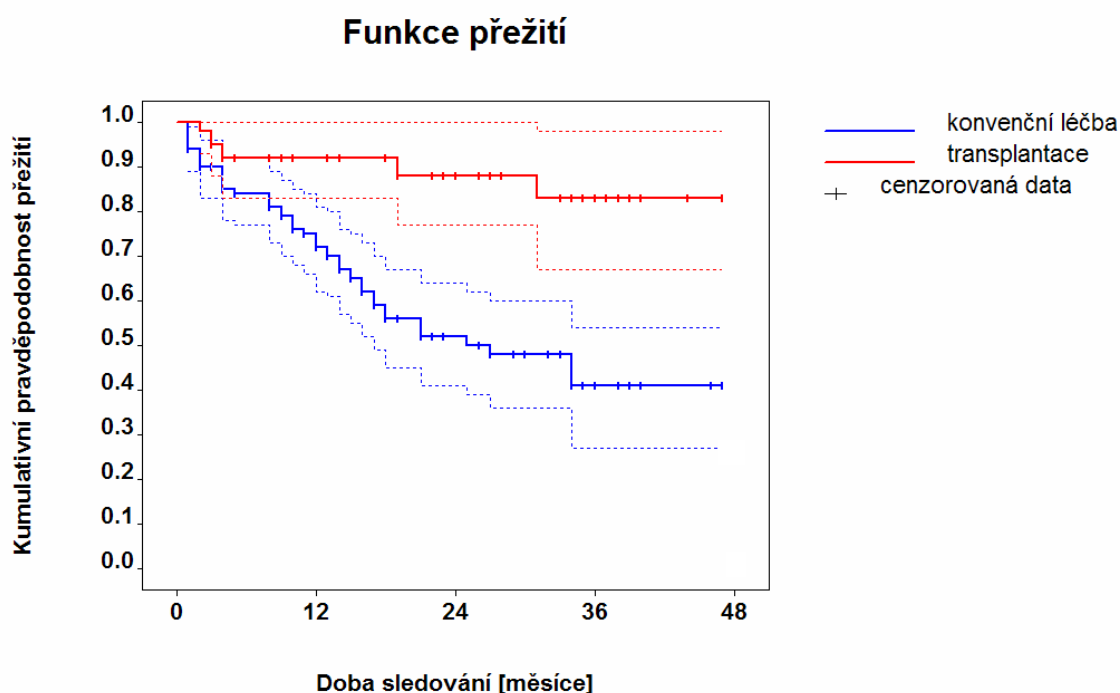
a) Vliv léčby na přežívání, Kaplan-Meierova analýza

**Case Processing Summary**

skupina	Total N	N of Events	Censored	
			N	Percent
konvenční léčba	101	44	57	56,4%
transplantovaní	37	5	32	86,5%
Overall	138	49	89	64,5%

Ve skupině konvenčně léčených chemoterapií bylo 101 pacientů, 44 pacientů (43,6 %) zemřelo, ve druhé skupině pacientů, kteří podstoupili autologní transplantaci kostní dřeně bylo 37 pacientů, 5 pacientů (13,5 %) zemřelo.

Kaplan-Meierovy křivky včetně 95% CI:



Odhad pravděpodobnosti přežití včetně 95% CI v zadaných časech:

Čas	konvenční léčba	transplantace
1 rok	0.747 (0.659-0.835)	0.917 (0.828-1.000)
3 roky	0.400 (0.263-0.538)	0.820 (0.667-0.974)

Means and Medians for Survival Time

skupina	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
konvenční léčba	27,355	2,076	23,287	31,423	25,000	6,067	13,109	36,891
transplantovaní	41,354	2,344	36,760	45,949	.	.	.	.
Overall	31,293	1,731	27,900	34,686	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Průměrná doba přežívání byla u skupiny konvenčně léčených 27,4 měsíců (95% CI 23,287 – 31,4 měs.) a u skupiny transplantovaných 41,4 měsíců (95% CI 36,8 – 45,9 měs.).

Hodnota mediánu přežívání byla odhadnuta pouze u skupiny konvenčně léčených a činila 25 měsíců (95% CI 13,1 – 36,9 měs.).

Porovnání křivek přežití log-rank testem:

#### Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	11,749	1	<b>,001</b>

Test of equality of survival distributions for the different levels of skup.

Hodnota signifikance log-rank testu  $p = 0,001$  ukazuje, že transplantovaní pacienti lépe přežívají.

b) Vliv hladiny beta2-mikroglobulinu na přežívání, Kaplan-Meierova analýza

Nejdříve byl vypočítán medián hodnot beta2-mikroglobulinu u všech pacientů, poté byli pacienti rozděleni na skupinu s nižšími hodnotami beta2-mikroglobulinu (nižšími než medián) a skupinu s vyššími hodnotami beta2-mikroglobulinu (vyššími než medián).

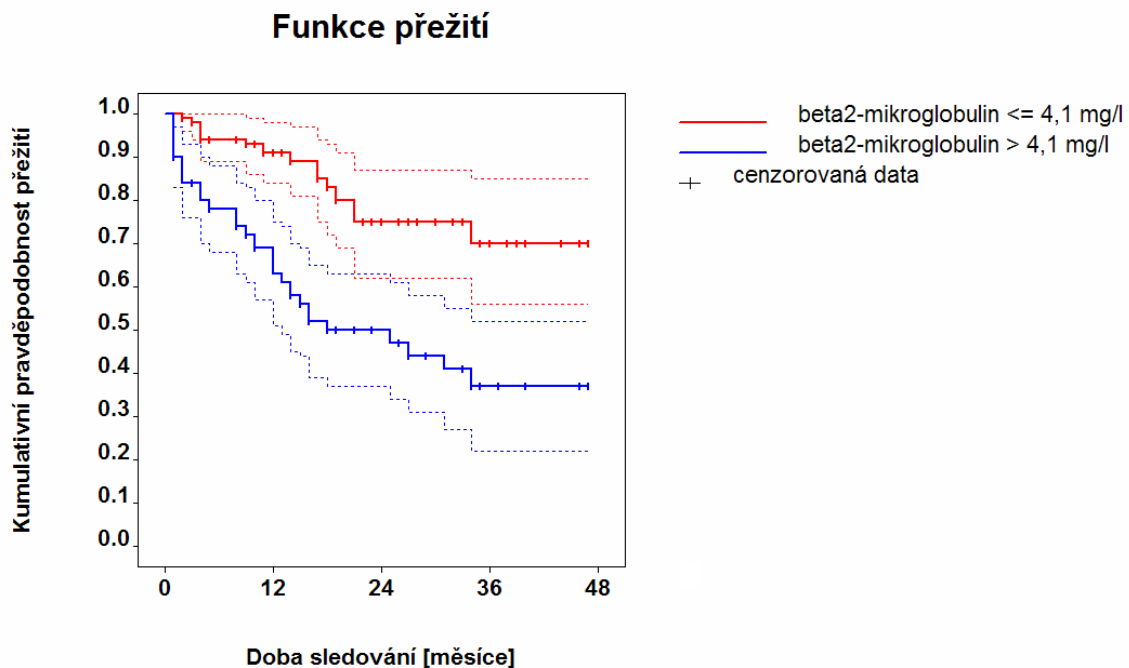
Hodnota mediánu pro celý soubor činila 4,1 mg/l.

**Case Processing Summary**

beta2-mikroglobulin	Total N	N of Events	Censored	
			N	Percent
< 4,1 mg/l	69	14	55	79,7%
> = 4,1 mg/l	69	35	34	49,3%
Overall	138	49	89	64,5%

Ve skupině pacientů s nižší hladinou beta2-mikroglobulinu (nižší než 4,1 mg/l) bylo 69 pacientů, 14 pacientů (20,3 %) zemřelo, ve druhé skupině pacientů s vyšší hladinou beta2-mikroglobulinu (vyšší než 4,1 mg/l), bylo také 69 pacientů, 35 pacientů (50,7 %) zemřelo.

Kaplan-Meierovy křivky včetně 95% CI:



Odhad pravděpodobnosti přežití včetně 95% CI v zadaných časech:

Čas	beta2-mikroglobulin ≤ 4,1 mg/l	beta2-mikroglobulin > 4,1 mg/l
1 rok	0.905 (0.833-0.978)	0.681 (0.568-0.795)
3 roky	0.697 (0.551-0.844)	0.363 (0.215-0.511)

Means and Medians for Survival Time

beta2-mikroglobulin	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
< 4,1 mg/l	37,865	2,111	33,728	42,002	.	.	.	.
> = 4,1 mg/l	24,995	2,473	20,148	29,842	18,000	6,820	4,633	31,367
Overall	31,293	1,731	27,900	34,686	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Průměrná doba přežívání byla u skupiny s nižší hladinou beta2-mikroglobulinu 37,9 měsíců (95% CI 33,7 – 42,0 měs.) a u skupiny s vyšší hladinou beta2-mikroglobulinu 25,0 měsíců (95% CI 20,1 – 29,8 měs.).

Hodnota mediánu byla odhadnuta pouze u skupiny s vyšší hladinou beta2-mikroglobulinu a činila 18 měsíců (95% CI 4,6 – 31,4 měs.).

#### Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	15,011	1	<b>,0001</b>

Test of equality of survival distributions for the different levels of b2m .

Hodnota signifikance log-rank testu  $p = 0,0001$  ukazuje, že pacienti s nižší hladinou beta2-mikroglobulinu lépe přežívají.

c) Vliv úrovně faktoru ICTP na přežívání, Kaplan-Meierova analýza

Nejdříve byl vypočítán medián hodnot faktoru remodelace kostní tkáně ICTP u všech pacientů, poté byli pacienti rozděleni na skupinu s nižšími hodnotami ICTP (nižšími než medián) a skupinu s vyššími hodnotami ICTP (vyššími než medián).

Hodnota mediánu pro celý soubor činila 8,3 ug/l.

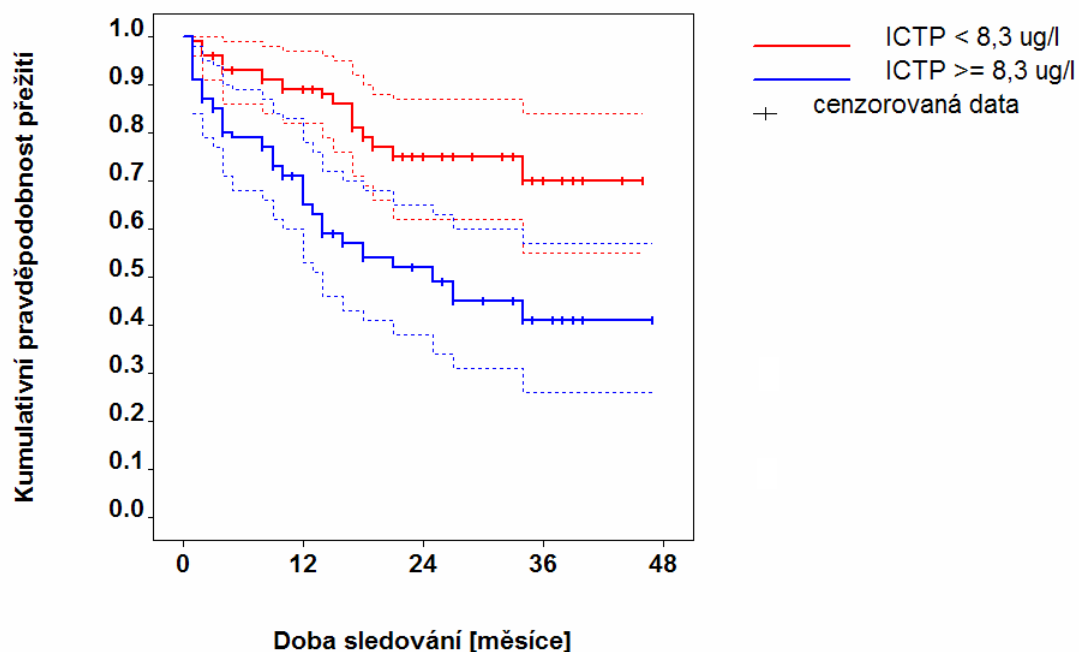
**Case Processing Summary**

ICTP	Total N	N of Events	Censored	
			N	Percent
< 8,3 ug/l	66	15	51	77,3%
>= 8,3 ug/l	66	30	36	54,5%
Overall	132	45	87	65,9%

Ve skupině pacientů s nižší hladinou ICTP (nižší než 8,3 mg/l) bylo 66 pacientů, 15 pacientů (22,7 %) zemřelo, ve druhé skupině pacientů s vyšší hladinou ICTP (vyšší než 8,3 mg/l), bylo také 66 pacientů, 30 pacientů (45,5 %) zemřelo.

Kaplan-Meierovy křivky včetně 95% CI:

### Funkce přežití





Odhad pravděpodobnosti přežití včetně 95% CI v zadaných časech:

Čas	ICTP < 8,3 ug/l	ICTP >= 8,3 ug/l
1 rok	0.889 (0.811-0.967)	0.709 (0.594-0.823)
3 roky	0.691 (0.544-0.837)	0.408 (0.254-0.561)

#### Means and Medians for Survival Time

ICTP	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
< 8,3 ug/l	36,561	2,102	32,442	40,681	.	.	.	.
>= 8,3 ug/l	26,300	2,622	21,160	31,440	25,000	6,997	11,286	38,714
Overall	31,855	1,771	28,383	35,326	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Průměrná doba přežívání byla u skupiny s nižší hladinou ICTP 36,6 měsíců (95% CI 32,4 – 40,7 měs.) a u skupiny s vyšší hladinou ICTP 26,3 měsíců (95% CI 21,2 – 31,4 měs.).

Hodnota mediánu byla odhadnuta pouze u skupiny s vyšší hladinou ICTP a činila 25 měsíců (95% CI 11,3 – 38,7 měs.).

#### Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	10,183	1	<b>,001</b>

Test of equality of survival distributions for the different levels of ICTP.

Hodnota signifikance log-rank testu  $p = 0,001$  ukazuje, že pacienti s nižší hladinou hodnot faktoru remodelace kostní tkáně ICTP lépe přežívají.

### **7.3 Studie sledování průchodnosti stehenní tepny po subintimální angioplastice**

V klinické studii byla sledována skupina 115 pacientů s uzavřenou stehenní tepnou. Pacientům byla na Radiologické klinice LF a FN Olomouc provedena subintimální angioplastika, při níž byla zprůchodněna tepna. Poté byli pacienti v pravidelných intervalech sledováni a zjišťovalo se, zda je tepna stále průchodná. Klinická studie si kladla za cíl zjistit, zda má vliv délka uzávěru a kvalita krevního řečiště pod uzávěrem, která je dána počtem průchozích tepen pod uzávěrem, na dobu průchodnosti tepny po zákroku. U pacientů byly sledovány další faktory viz. tabulka 7.3.1 a bylo zjišťováno, zda jsou významnými prediktory pro čas průchodnosti tepny od zákroku.

Tabulka 7.3.1 Vstupní data

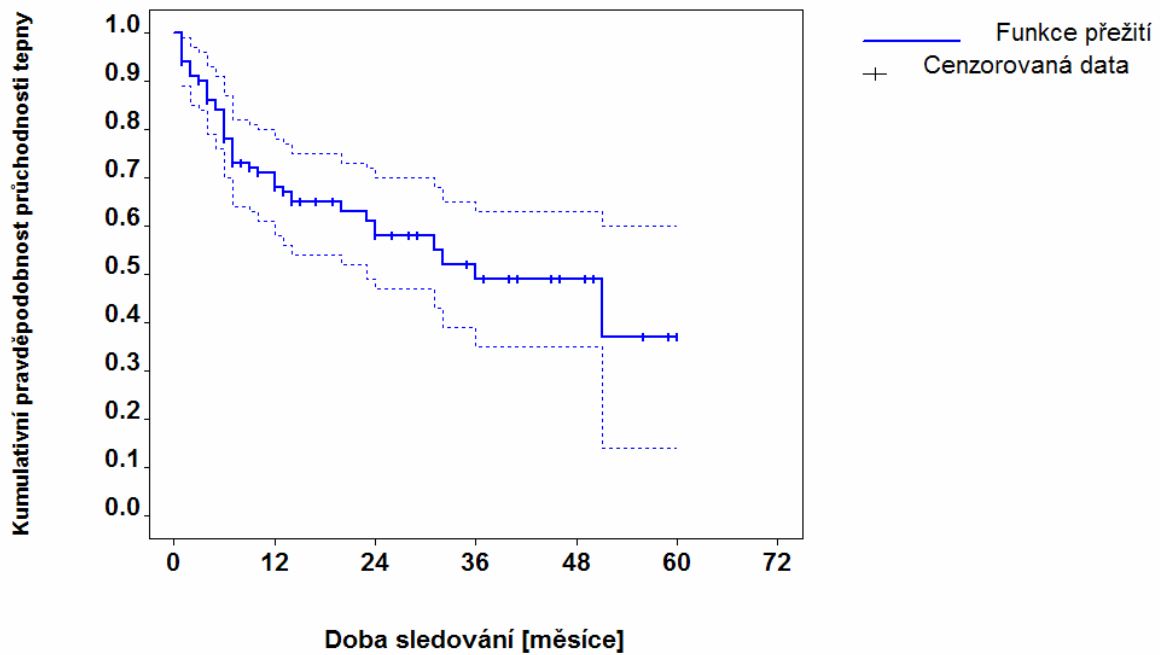
Číslo pacienta	Pohlaví		Lokalizace uzávěru		Indikace		Kvalita výtokového traktu		Délka uzávěru tepny [cm]		ICHS	DM	Hyperlipidémie	Kouření	CHOPN	Renální insuficience	Hypertenze	IM	CMP	Doba sledování [měsíce]	Průchodnost tepny	
	0	1	0	1	0	1	0	1	0	1												0
1	0	2	0	1	25	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	
2	1	4	1	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	
3	0	2	1	0	30	0	1	0	0	0	0	0	0	0	0	0	1	0	0	1	1	
4	0	3	1	0	4	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	
5	1	1	1	0	12	1	1	1	1	1	1	0	0	0	0	0	1	0	0	1	1	
6	1	4	1	0	15	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	
7	1	1	1	1	20	1	0	1	0	0	0	0	0	0	0	0	0	1	0	1	1	
8	0	2	1	0	7	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2	1	
9	0	2	1	1	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	
10	1	2	0	1	12	1	0	1	0	0	0	0	0	0	0	0	0	0	0	2	1	
11	0	2	0	1	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	
12	0	2	1	0	20	1	1	0	0	0	0	0	0	0	0	0	1	0	0	3	0	
13	1	1	0	1	7	0	1	0	0	0	0	0	0	0	0	0	1	0	0	4	1	
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
109	0	1	0	1	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	49	0	
110	0	1	0	1	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	50	0	
111	1	1	0	1	25	0	0	0	0	0	0	0	0	0	0	0	1	0	0	50	0	
112	1	1	0	0	2	1	0	0	1	0	0	0	0	0	0	0	1	0	0	51	1	
113	1	1	0	1	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	56	0	
114	0	1	0	1	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	59	0	
115	1	1	0	1	10	0	1	0	1	0	0	0	0	0	0	0	0	0	0	60	0	

**Vysvětlivky:**

Pohlaví 0 – žena, 1 – muž  
 Lokalizace uzávěru 1 - stehenní tepna, 2 – podkoleenní tepna, 3 – uzávěr přes obě tepny, 4 - uzávěr uzávěr přes obě tepny a trifurkaci  
 Indikace 0 – klaudikace (pacienti pocítují pouze bolest při chůzi, 2 – chronická kritická končetinová ischemie  
 Kvalita výtokového traktu (dána počtem průchozích tepen v krevním řečišti pod uzávěrem) 0 – 0 až 1 průchozí tepna, 1 – 2-3 průchozí tepny  
 ICHS (ischemická choroba srdeční) 0 – ne, 1 - ano  
 DM (diabetes mellitus) 0 – ne, 1 - ano  
 Hyperlipidémie 0 – ne, 1 - ano  
 Kouření 0 – ne, 1 - ano  
 CHOPN (chronická obstrukční plicní nemoc) 0 – ne, 1 - ano  
 Renální insuficience - 0 – ne, 1 - ano  
 Hypertenze 0 – ne, 1 - ano  
 IM (infarkt myokardu) 0 – ne, 1 - ano  
 CMP (cévní mozková příhoda) 0 – ne, 1 - ano  
 Průchodnost tepny 0 – tepna průchodná, 1 – došlo k zneprůchodnění

Kaplan-Meierova funkce přežití včetně 95% CI má tvar:

### Funkce přežití



Odhad pravděpodobnosti přežití v zadaných časech:

Čas	Pravděpodobnost průchodnosti tepny (95% CI)
1 rok	0.702 (0.607-0.797)
3 roky	0.519 (0.388-0.649)
5 let	0.363 (0.133-0.593)

Coxova regrese – počítačové zpracování:

**Case Processing Summary**

		N	Percent
Cases available in analysis	Even <sup>a</sup>	37	32,2%
	Censored	56	48,7%
	Total	93	80,9%
Cases dropped	Cases with missing values	4	3,5%
	Cases with negative time	0	,0%
	Censored cases before the earliest event in a stratum	18	15,7%
	Total	22	19,1%
Total		115	100,0%

a. Dependent Variable: Doba sledování [měsíce]

Tabulka udává přehled 93 případů (pacientů), z nichž 56 časových údajů bylo cenzorováno a u 37 pacientů došlo k události (k zneprůchodnění tepny). 22 případů bylo vypuštěno, u 4 případů nebyly kompletní údaje a 18 případů bylo cenzorováno dříve než došlo k události u prvního pacienta.

**Variables not in the Equation<sup>a</sup>**

	Score	df	Sig.
pohlaví	2,263	1	,132
Lokalizace uzávěru	6,596	3	,086
Lokalizace uzávěru(1)	,601	1	,438
Lokalizace uzávěru(2)	,711	1	,399
Lokalizace uzávěru(3)	5,335	1	<b>,021</b>
Indikace	2,155	1	,142
Výtok.trakt	6,592	1	<b>,010</b>
ICHS	,059	1	,808
DM	,108	1	,743
Hyperlipidémie	,208	1	,649
Kouření	,048	1	,826
CHOPN	2,447	1	,118
Ren.insuf	,968	1	,325
Hypertenze	,042	1	,837
IM	,447	1	,504
CMP	,057	1	,811
Délka uzávěru cm	8,009	1	<b>,005</b>

a. Residual Chi Square = 27,119 with 16 df Sig. = ,040

Tabulka uvádí všechny vstupní parametry s vypočítanými statistickými signifikancemi. Je zřejmé, že k významným prediktorům zneprůchodnění tepny bude patřit kvalita výtokového traktu a délka uzávěru.

Variables in the Equation

		B	SE	Wald	df	Sig.	Exp(B)	95,0% CI for Exp(B)	
								Lower	Upper
Step 1	Délka uzávěru cm	,055	,020	7,688	1	,006	1,056	1,016	1,098
Step 2	Výtok.trakt	-1,210	,394	9,435	1	<b>,002</b>	,298	,138	,645
	Délka uzávěru cm	,066	,020	10,763	1	<b>,001</b>	1,068	1,027	1,112

Metodou postupného přidávání nezávisle proměnných do modelu byly ve dvou krocích postupně navrženy dva modely. U proměnných zařazených do modelu byly opět odhadnuty koeficienty B, jejich chyby SE, hodnoty Waldovy statistiky (Wald), počet stupňů volnosti (df), signifikance (Sig.) a bylo spočítáno Exp(B), tedy hazard ratio HR (změna rizika exitu při jednotkové změně nezávisle proměnné) a 95% intervaly spolehlivosti pro HR.

Významné faktory pro predikci zneprůchodnění uzávěru tepny jsou:

- Kvalita výtokového traktu – více průchodných bércoých tepen ve výtokovém traktu (2 – 3) snižuje riziko opětovného zneprůchodnění stehenní tepny 0,298x (na 29,8 % původního rizika). 95% CI je 0,138 - 0,645.
- Délka uzávěru stehenní tepny - zvětší-li se délka uzávěru stehenní tepny o 1 cm, zvyšuje se riziko opětovného zneprůchodnění stehenní tepny 1,068x (o 6,8 %). 95% CI je 1,027 – 1,112.

Na základě těchto poznatků byla provedena Kaplan-Meierova analýza pro pacienty rozdělené dle kvality výtokového traktu a dle délky uzávěru stehenní tepny:

a) Kaplan-Meierovy křivky pro 2 skupiny pacientů rozdělené podle kvality výtokového traktu.

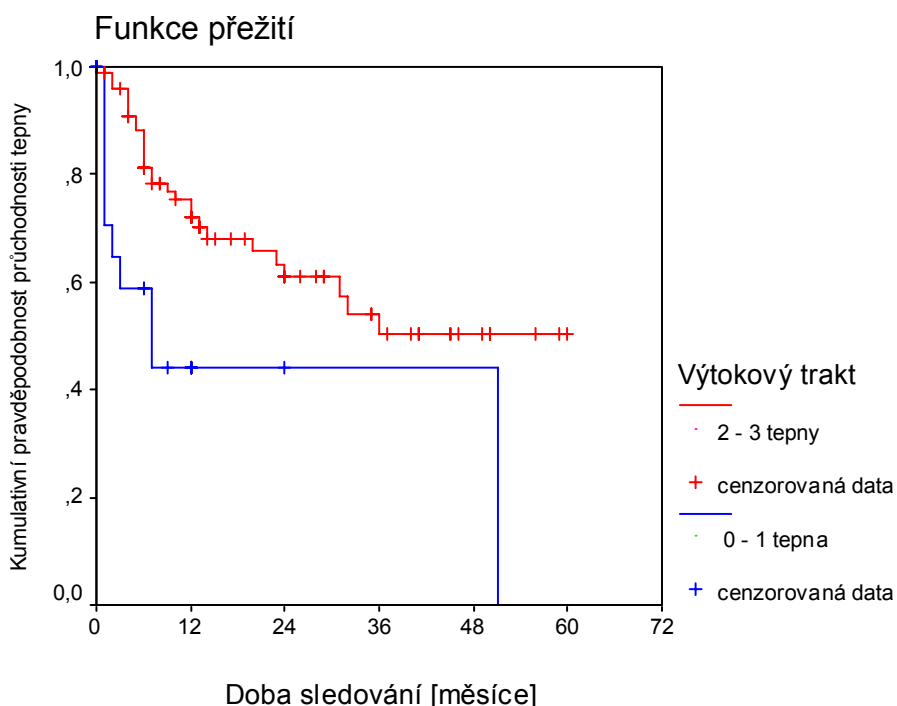
První skupina (v tabulkách značena výtok. trakt horší) má v krevním řečišti pod uzávěrem stehenní tepny průchodnou pouze jednu nebo žádnou bér covou tepnu, druhá skupina (v tabulkách značena výtok. trakt lepší) má průchodné 2 – 3 bér cové tepny.

**Case Processing Summary**

Výtok.trakt	Total N	N of Events	Censored	
			N	Percent
horší	21	10	11	52,4%
lepší	93	28	65	69,9%
Overall	114	38	76	66,7%

Ve skupině s horší kvalitou výtokového traktu bylo 21 pacientů, u 10 (47,6 %) došlo k zneprůchodnění tepny, ve druhé skupině s lepší kvalitou výtokového traktu bylo 93 pacientů, k zneprůchodnění tepny došlo u 28 (30,1%).

Kaplan-Meierovy křivky:



### Means and Medians for Survival Time

Výtok.trakt	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
horší	24,118	6,404	11,566	36,670	7,000	3,454	,230	13,770
lepší	37,533	3,235	31,191	43,874	.	.	.	.
Overall	34,372	3,031	28,431	40,312	36,000	8,895	18,567	53,433

a. Estimation is limited to the largest survival time if it is censored.

Průměrná doba do opětovného zneprůchodnění stehenní tepny byla u skupiny s horší kvalitou výtokového traktu 24,1 měsíců (95% CI 11,6 – 36,7 měs.) a u skupiny s lepší kvalitou 37,5 měsíce (95% CI 31,2 – 43,9 měs.).

Hodnota mediánu byla odhadnuta pouze u skupiny s horší kvalitou výtokového traktu a činila 7 měsíců (95% CI 0,23 – 13,77 měs.).

Porovnání křivek přežití log-rank testem:

#### Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	7,264	1	<b>,007</b>

Test of equality of survival distributions for the different levels of výtok.trakt.

Hodnota signifikance log-rank testu  $p = 0,007$  ukazuje, že pacienti s kvalitnějším krevním řečištěm (s více průchodnými bérčovými tepnami) mají vyšší pravděpodobnost, že po provedeném zákroku zůstane stehenní tepna průchozí.

b) Kaplan-Meierovy křivky pro 2 skupiny pacientů rozdělené podle délky uzávěru stehenní tepny.

První skupina má délku uzávěru menší nebo rovnu 15 cm, druhá skupina má délku uzávěru větší než 15 cm.

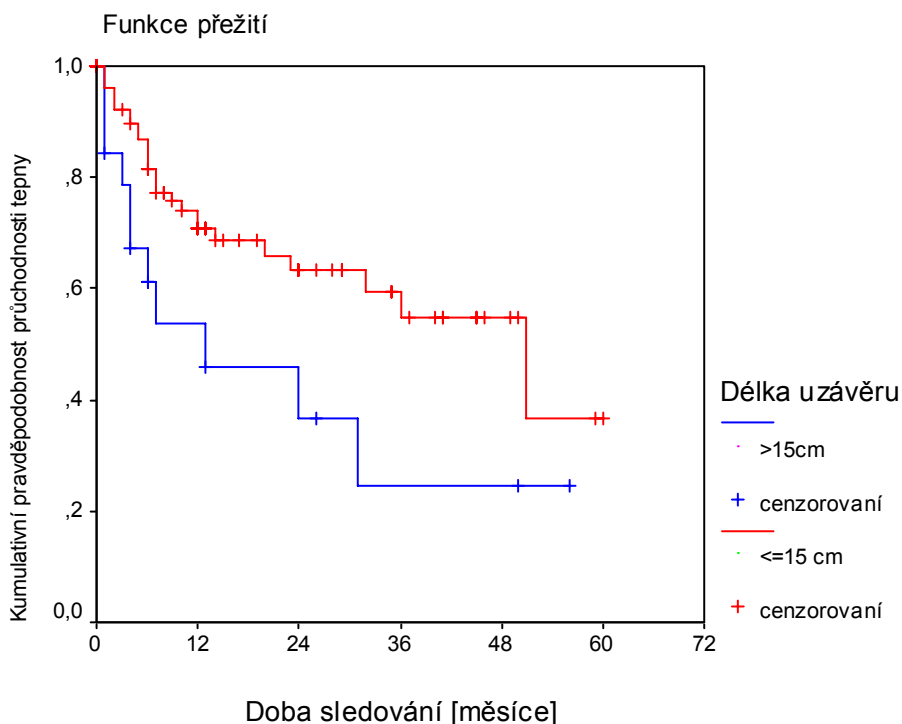
#### Case Processing Summary

Délka uzávěru	Total N	N of Events	Censored	
			N	Percent
<=15 cm	93	27	66	71,0%
>15 cm	22	11	11	50,0%
Overall	115	38	77	67,0%



Ve skupině s délkou uzávěru menší nebo rovnou 15 cm bylo 93 pacientů, u 27 (29,0 %) došlo k zneprůchodnění tepny, ve druhé skupině s délkou uzávěru větší než 15 cm bylo 22 pacientů, k zneprůchodnění tepny došlo u 11 (50,0 %).

Kaplan-Meierovy křivky:



Means and Medians for Survival Time

Délka uzávěru	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
<=15 cm	37,244	3,359	30,661	43,827	51,000	13,159	25,209	76,791
>15cm	22,395	5,754	11,116	33,673	13,000	9,383	,000	31,391
Overall	34,632	3,017	28,719	40,545	36,000	8,795	18,762	53,238

a. Estimation is limited to the largest survival time if it is censored.

Průměrná doba do opětovného zneprůchodnění stehenní tepny byla u skupiny s kratším uzávěrem stehenní tepny 37,2 měsíců (95% CI 30,7 – 43,8 měs.) a u skupiny s delším uzávěrem stehenní tepny 22,4 měsíců (95% CI 11,1 – 33,7 měs.).

Hodnota mediánu u skupiny s kratším uzávěrem stehenní tepny činila 51 měsíců (95% CI 25,2 – 76,8 měs.) a u skupiny s delším uzávěrem stehenní tepny 13 měsíců (95% CI 0,00 – 31,4 měs.)

Porovnání Kaplan-Meierových křivek log-rank testem:

**Overall Comparisons**

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	5,110	1	<b>,024</b>

Test of equality of survival distributions for the different levels of  
Délka uzávěru.

Hodnota signifikance log-rank testu  $p = 0,024$  ukazuje, že pacienti s uzávěrem stehenní tepny menším nebo rovným 15 cm mají vyšší pravděpodobnost, že po provedeném zákroku zůstane stehenní tepna průchozí než pacienti s delším uzávěrem.

## **7.4 Studie sledování přežívání endoprotéz u pacientů po operaci kyčelního kloubu**

V klinické studii, která byla provedena ve spolupráci s Ortopedickou klinikou LF a FN v Olomouci, byla sledována skupina 127 pacientů po operaci kyčelního kloubu. Celkem bylo u těchto pacientů nahrazeno totální endoprotézou 137 kyčelních kloubů. Po operaci byli pacienti zváni na kontroly, které spočívaly v klinickém a rentgenologickém vyšetření. Pokud endoprotéza selhala, byla celá nebo její část vyměněna. Tento zákrok je nazýván revizí. U pacientů byly sledovány další faktory (viz. tabulka 7.4.1) a bylo zjišťováno, které z těchto faktorů mají vliv na přežívání totálních endoprotéz.

Tabulka 7.4.1 Vstupní data

Číslo pacienta	Pohlaví	Věk v době operace [roky]	Předoperační HHS	Strana	Velikost jamky [mm]	Typ polyetylénu	Revize	Délka přežití endoprotézy [roky]
1	m	44	46	dx	58	SI	0	12,2
2	m	50	56	sin	54	SI	0	12,4
3	ž	42	62	dx	52	SI	1	6,5
4	ž	45	46	sin	52	SI	1	7,6
5	ž	56	46	dx	50	SI	0	12,3
6	ž	43	57	dx	52	SI	0	12,2
7	ž	43	48	sin	52	SI	0	11,5
8	m	42	56	dx	60	SI	1	10,4
9	m	41	46	sin	54	SI	0	12,2
10	ž	50	56	dx	50	SI	1	12,3
11	ž	39	62	sin	52	A	0	12,4
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
131	ž	48	56	sin	54	SI	0	11,1
132	ž	60	56	dx	52	A	0	11,1
133	m	46	56	dx	60	SI	0	10,5
134	m	44	46	dx	50	SI	0	10,5
135	ž	49	48	sin	54	SI	1	9,2
136	m	43	54	dx	54	SI	0	11,1
137	ž	50	48	dx	50	SI	0	10,5

Vysvětlivky

Věk v době operace – věk pacientů v letech v době, kdy jim byla voperována endoprotéza

Pohlaví m – muž, ž - žena

Předoperační HHS – Harrisovo skóre zjištěné před operací, HHS se používá k hodnocení výsledku operace náhrady kyčelního kloubu, hodnotí např. míru bolesti a funkčnost kloubu (zda pacient kulhá, kolik ujde metrů atd.)

Strana – na které straně byla kyčel operována, dx – vpravo, sin – vlevo

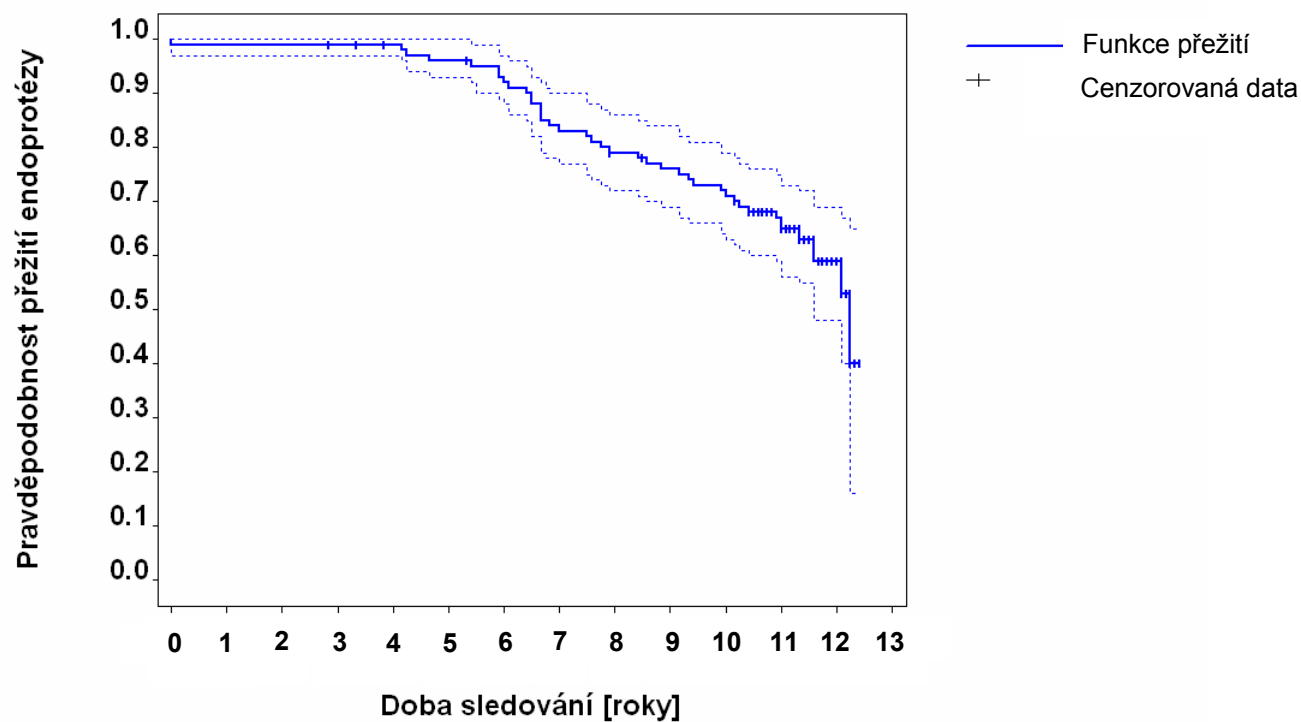
Velikost jamky – velikost implantátu použitého při první operaci v mm

Typ polyetylénu – typ polyetylénu v jamce použitého implantátu SI – standardní vložka, A – antiluxační vložka

Revize 1 – došlo k revizi operované kyčle, 0 – nedošlo k revizi

Kaplan-Meierova funkce přežití má tvar:

### Funkce přežití



Odhad pravděpodobnosti přežití v zadaných časech:

Čas	Pravděpodobnost přežití (95% CI)
1 rok	0.985 (0.965-1.000)
3 roky	0.985 (0.965-1.000)
5 let	0.955 (0.920-0.990)
7 let	0.834 (0.771-0.897)
10 let	0.711 (0.634-0.789)

Coxova regrese – počítačové zpracování:

#### Case Processing Summary

		N	Percent
Cases available in analysis	Event <sup>a</sup>	51	37,2%
	Censored	86	62,8%
	Total	137	100,0%
Cases dropped	Cases with missing values	0	,0%
	Cases with negative time	0	,0%
	Censored cases before the earliest event in a stratum	0	,0%
	Total	0	,0%
Total		137	100,0%

a. Dependent Variable: Délka přežití v letech

Tabulka udává přehled 137 operovaných kyčelních kloubů, z nichž 86 dat bylo cenzorováno a u 51 kloubů došlo k revizi. Žádný případ nebyl vypuštěn.

#### Variables not in the Equation<sup>a</sup>

	Score	df	Sig.
Pohlaví	3,033	1	,082
Věk v době operace	6,032	1	<b>,014</b>
Předoperační HHS	,805	1	,370
Strana	,079	1	,778
Velikost jamky v mm	8,704	1	<b>,003</b>
Typ polyetylénu	,000	1	,989

a. Residual Chi Square = 16,590 with 6 df Sig. = ,011

Tabulka uvádí všechny vstupní parametry s vypočítanými statistickými signifikancemi. Je zřejmé, že k významným prediktorům revize bude patřit věk pacientů v době operace a velikost jamky implantátu.

#### Variables in the Equation

		B	SE	Wald	df	Sig.	Exp(B)	95,0% CI for Exp(B)	
								Lower	Upper
Step 1	Velikost jamky v mm	-,143	,049	8,534	1	,003	,867	,788	,954
Step 2	Věk v době operace	-,049	,021	5,389	1	<b>,020</b>	,953	,914	,992
	Velikost jamky v mm	-,137	,048	8,054	1	<b>,005</b>	,872	,793	,958

Metodou postupného přidávání nezávisle proměnných do modelu byly ve dvou krocích postupně navrženy dva modely. U proměnných zařazených do modelu byly opět odhadnuty koeficienty B, jejich chyby SE, hodnoty Waldovy statistiky (Wald), počet stupňů volnosti

(df), signifikance (Sig.) a bylo spočítáno Exp(B), tedy hazard ratio HR (změna rizika exitu při jednotkové změně nezávisle proměnné) a 95% intervaly spolehlivosti pro HR.

Významné faktory pro predikci revize kyčelního kloubu jsou:

- věk pacienta v době operace – zvýší-li se věk pacienta o jeden rok, sníží se riziko revize operované kyčle 0,953x (o 4,7 %). 95% CI je 0,914-0,992.
- velikost jamky implantátu – zvětší-li se velikost jamky o 1 cm, sníží se riziko revize operované kyčle 0,872x (o 12,8 %). 95% CI je 0,793-0,958.

Na základě těchto poznatků byla provedena Kaplan-Meierova analýza pro pacienty rozdělené dle věku a velikosti jamky:

a) Kaplan-Meierovy křivky přežití pro 2 skupiny operovaných kyčelních kloubů, které jsou rozděleny podle věku pacientů.

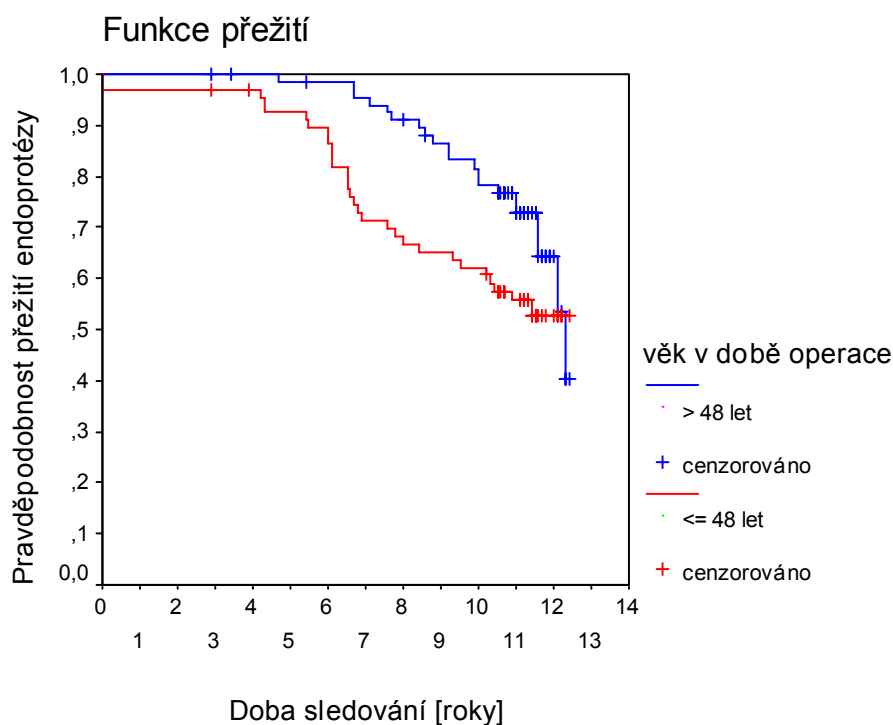
Byla spočítána hodnota mediánu pro veličinu věk a podle ní byli operované kyčelní klouby rozděleny do dvou skupin, mladší skupina (věk  $\leq 48$  let) a starší skupina (věk  $> 48$  let).

**Case Processing Summary**

Věk v době operace	Total N	N of Events	Censored	
			N	Percent
$\leq 48$ let	68	30	38	55,9%
$> 48$ let	69	21	48	69,6%
Overall	137	51	86	62,8%

Ve skupině operovaných kyčelních kloubů u mladších pacientů bylo 68 kloubů, u 30 (44,1 %) došlo k revizi, ve druhé skupině operovaných kyčlí u starších pacientů bylo 69 kloubů, k revizi došlo pouze u 21 (30,4 %) kyčlí.

Kaplan-Meierovy křivky:



**Means and Medians for Survival Time**

Věk v době operace	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
<=48 let	9,868	,406	9,073	10,664	.	.	.	.
>48 let	11,286	,226	10,843	11,729	12,315	,394	11,543	13,087
Overall	10,586	,241	10,114	11,058	12,315	,461	11,411	13,219

a. Estimation is limited to the largest survival time if it is censored.

Průměrná doba přežívání totálních endoprotéz byla u skupiny operovaných kyčelních kloubů u mladších pacientů 9,87 roků (95% CI 9,07 – 10,66 roků) a u skupiny operovaných kyčelních kloubů u starších pacientů 11,28 roků (95% CI 10,84 – 11,73 roků).

Hodnota mediánu byla spočítána pouze u skupiny operovaných kyčlí u starších pacientů a činila 12,32 roků (95% CI 11,54 – 13,09 roků).

Porovnání křivek přežití log-rank testem:

**Overall Comparisons**

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	4,327	1	<b>,038</b>

Test of equality of survival distributions for the different levels of Věk v době operace.



Hodnota signifikance log-rank testu  $p = 0,038$  prokazuje, že totální endoprotézy voperované starším pacientům mají lepší přežívání než u mladších pacientů.

b) Kaplan-Meierovy křivky přežití pro 2 skupiny operovaných kyčelních kloubů, které byly rozděleny podle velikosti jamky implantátu.

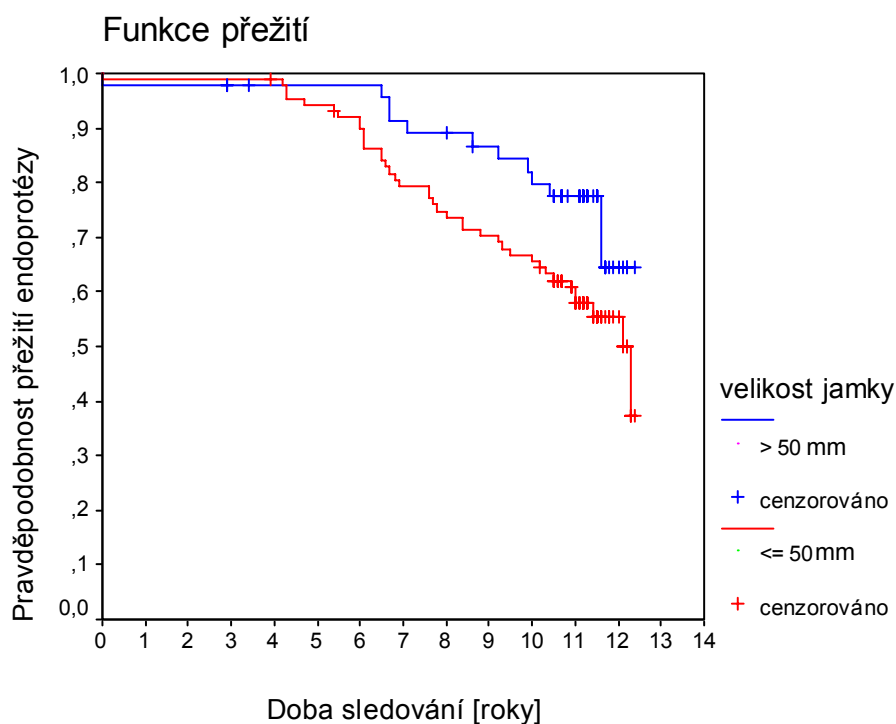
První skupina má velikost jamky implantátu menší nebo rovnou 50 mm, druhá skupina má velikost jamky implantátu větší než 50 mm.

**Case Processing Summary**

Velikost jamky	Total N	N of Events	Censored	
			N	Percent
< = 50 mm	56	29	27	48,2%
> 50 mm	81	22	59	72,8%
Overall	137	51	86	62,8%

Ve skupině operovaných kyčelních kloubů s velikostí jamky implantátu menší nebo rovnou 50 mm bylo 56 kyčlí, u 29 (51,8 %) došlo k revizi, ve druhé skupině operovaných kyčelních kloubů s velikostí jamky implantátu větší než 50 mm bylo 81 kyčlí, k revizi došlo pouze u 22 (27,2 %) kyčlí.

Kaplan-Meierovy křivky:



### Means and Medians for Survival Time

Velikost jamky	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
<= 50 mm	9,822	,411	9,018	10,627	11,400	,655	10,116	12,684
> 50 mm	11,101	,276	10,561	11,641	.	.	.	.
Overall	10,574	,239	10,104	11,043	12,300	,508	11,305	13,295

a. Estimation is limited to the largest survival time if it is censored.

Průměrná doba přežívání totálních endoprotéz byla u skupiny operovaných kyčelních kloubů s menší velikostí jamky implantátu 9,82 roků (95% CI 9,02 – 10,63 roků) a u skupiny operovaných kyčlí s větší velikostí jamky implantátu 11,10 roků (95% CI 10,56 – 11,64 roků).

Hodnota mediánu byla spočítána pouze u skupiny operovaných kyčelních kloubů s menší velikostí jamky implantátu a činila 11,40 roků (95% CI 10,11 – 12,68 roků).

Porovnání křivek přežití log-rank testem:

#### Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	8,084	1	<b>,004</b>

Test of equality of survival distributions for the different levels of velikost jamky.

Hodnota signifikance log-rank testu  $p = 0,004$  prokazuje, že totální endoprotézy s větší jamkou mají lepší přežívání než endoprotézy s menší jamkou.

## 8 Závěr

V rámci této práce byl podán přehledný souhrn metod, které používá analýza přežívání pro zkoumání cenzorovaných dat. Tyto metody usilují vytěžit maximum relevantní informace v případech, kdy část dat nebyla ve chvíli ukončování sběru dat úplná; existuje pro ně jen „dolní odhad“. Jedná se o situace, které můžeme zahrnout do třídy „přežití – survival“. Klasickou situací dat tohoto typu jsou údaje přežití pro pacienty s nádorovým onemocněním. V praktické části byla navržena a implementována aplikace, která doplňuje možnosti statistického programu SPSS o dvě funkce. První funkce umožňuje zobrazit jednu nebo více křivek přežití včetně 95% intervalu spolehlivosti. Druhá funkce spočívá v možnosti určit z uvedených křivek přežití pravděpodobnost přežití včetně intervalů spolehlivosti v libovolně zvolených časech a tato data zobrazit jako tabulku. Vznik této aplikace si vyžádaly požadavky uživatelů z řad lékařů.

Dále jsou metodami analýzy přežívání prakticky řešeny čtyři klinické studie. Data ke všem studiím poskytly kliniky Lékařské fakulty a Fakultní nemocnice. První sleduje přežívání hemato-onkologicky nemocných pacientů a faktory, které toto přežívání ovlivňují. Studie byla provedena ve spolupráci s Hemato-onkologickou klinikou.

Druhá studie sleduje obdobně přežívání pacientů s mnohočetnými myelomy a hledají se významné faktory ovlivňující toto přežívání. Tato studie vznikla ve spolupráci se III. interní klinikou.

Třetí studie řešená metodami analýzy přežívání se od prvních dvou liší tím, že sledovaným jevem není úmrtí pacientů, ale opětovné zneprůchodnění tepny u pacientů, kterým byla z důvodu uzavření stehenní tepny provedena subintimální angioplastika. Data byla získána od lékařů Radiologické kliniky.

Poslední, čtvrtá studie sleduje přežívání totálních endoprotéz u pacientů po operaci kyčelního kloubu. Sledovaným jevem je v této studii revize totální endoprotézy.

Závěrem je třeba zdůraznit nebezpečí špatného interpretování statistických výsledků v případě, kdy aplikujeme standardní statistické metody, např. výpočet aritmetického průměru a mediánu na cenzorovaná data.

## **Souhrn**

Mezi statistickými metodami používanými pro vyhodnocování dat biomedicínského výzkumu se zhruba od konce 50 let minulého století jako stále významnější a použitelnější prokazuje metoda pracující s datovými soubory, které kromě konečných dat mohou obsahovat i data cenzorovaná. K cenzorování dochází tehdy, pokud ve sledované době nenastane jistá událost, která je předmětem sledování. Metody pracující s cenzorovanými daty se souhrně označují analýza přežívání.

Tyto metody dnes nacházejí uplatnění i v dalších oborech např. ve strojírenství a elektrotechnice.

Statistické oddělení Ústavu lékařské biofyziky aplikuje metody analýzy přežití už zhruba dvacet let. Impulsem k tomu byla potřeba vyhodnocovat data přežívání především onkologických pacientů shromažďovaná lékaři internisty a pracovníky Hemato-onkologické kliniky.

V této disertační práci byly teoreticky studovány možnosti metod analýzy přežívání, tedy vlastnosti funkce přežití a hazardu, odhady distribuce přežívání, testy umožňující porovnat dvě a více křivek přežití a jsou zde objasněny základní principy Coxova regresního modelu.

Tyto teoretické poznatky jsou aplikovány na data čtyř klinických studií, které probíhaly na klinikách Lékařské fakulty UP a Fakultní nemocnice v Olomouci.

Pro snadnější a přesnější počítačové zpracování dat analýzy přežívání byla vyvinuta aplikace doplňující současně používaný software o dvě funkce. První funkce umožňuje zobrazit jednu nebo více křivek přežití včetně 95% intervalu spolehlivosti. Druhá funkce spočívá v možnosti určit z uvedených křivek přežití pravděpodobnost přežití včetně intervalů spolehlivosti v libovolně zvolených časech a tato data zobrazit jako tabulku.

## **Klíčová slova**

analýza přežívání, cenzorovaná data, Kaplan-Meierovy křivky přežití, Coxův regresní model

## **Summary**

Among the statistical methods applied to data evaluation in the biomedical research there is one that has been proving to be more and more important and significant since the middle of the last century. Censoring occurs when a given event which is a subject of observation will not happen in a given time period. Methods working with censored data are generally called survival analysis.

Nowadays these methods are also used in other branches eg. in mechanical and electrical engineering.

The statistical department of the Institute of Medical Biophysics has been applying survival analysis for more than two decades. The first impuls to start with the method was a requirement for an analysis of incomplete survival data obtained mostly from oncological patients and collected by internal physicians and hematooncologists.

In the Ph.D. thesis possibilities of survival analysis methods were theoretically studied. In particular, there are explained features of survival and hazard functions, estimations of survival distribution, tests to compare two or more survival curves and the basic principles of Cox regressions model. This theory is applied to data of four clinical studies performed on clinics of the University Hospital Olomouc.

To make the computer processing of survival analysis data easier and more precise a new software application was developed that adds two new functions to the currently used software. The first function enables to display one or more survival functions including its 95% confidence interval. The second function gives us the possibility to get the survival probability values including 95% confidence interval from the survival functions for arbitrarily given times and display these values in a table.

## **Keywords**

survival analysis, censored data, Kaplan-Meier survival curves, Cox regression model

## Literatura

- [1] Alling D. Early decision in the Wilcoxon two-sample test. *J. Amer. Statist. Ass.* 1963; 58: 713 – 20.
- [2] Anděl J. *Matematická statistika*. SNTL – Nakladatelství technické literatury, Praha, 1985.
- [3] Altman D. G. *Practical statistics for medical research*. Chapman & Hall, UK 1995.
- [4] Bencko V. *Statistické metody v epidemiologii*, Nakladatelství Karolinum, Praha, 2003
- [5] Breslow, N. E. Covariance analysis of censored survival data. *Biometrics*, 1974; 30: 89-99.
- [6] Bryan F. M. *Multivariate Statistical Methods – A Primer*. Chapman & Hall, UK 2004.
- [7] Cain, K. C., Lange T. Approximate case influence for the proportional hazards regression model with censored data. *Biometrics*, 1984; 40: 493-499.
- [8] Cox D.R., Oakes D. *Analysis of Survival Data*. Chapman & Hall, USA, 1998.
- [9] Cox D. R. Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B*, 1972; 34: 187-220.
- [10] Elandt-Johnson R. C., Johnson N. L.: *Survival Models and Data Analysis*. John Wiley & Sons, Inc., New York, 1980.
- [11] Galló J., Langová K., Havránek V., Čechová I.: Poor survival of ABG I hip prosthesis in younger patients. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub.* 2008; 152(1):163-169.
- [12] Gehan E. A. A generalized Wilcoxon test for comparing arbitrarily singly-censored samples, *Biometrika* 1965, 52: 203-223.

- [13] Gehan, E. A. Statistical methods for survival time studies. In: Cancer Therapy: Prognostic Factors and Criteria, M. J. Staquet, eds. New York: Raven Press; 1975.
- [14] Kalbfleisch, J. D., Prentice P.L. The statistical analysis of failure time data. New York: John Wiley & Sons, Inc., 1980.
- [15] Kaplan E. L., Meier P. Nonparametric estimation from incomplete observations. J Amer Statist Ass 1958; 53: 457 – 481.
- [16] Klein J. P., Goel P. K. Survival Analysis: State of the Art. Kluwer Academic Publisher, Dordrecht, The Netherlands, 1992.
- [17] Klein J. P., Moeschberger M. L. Survival Analysis: Techniques for Censored and Truncated Data. Springer, New York, USA, 1997.
- [18] Köcher M., Černá M., Utíkal P., Kozák J., Šišola I., Bachleda P., Dráč P., Sekanina Z., Langová K.: Subintimální angioplastika ve femoropopliteální oblasti – střednědobé výsledky. Ces Radiol 2008; 62(2): 153-159.
- [19] Langová K., Zapletalová J., Komenda S.: Statistická analýza přežívání. Sborník příspěvků XXXI. dny lékařské biofyziky, 2008: 86-92.
- [20] Lawless R. F. Statistical models and methods for lifetime data. John Wiley & Sons, Inc., New York , USA, 1982.
- [21] Lee E., Desu M. A computer program for comparing k samples with right censored data. Computer Programs in Biomedicine 1972; 2: 315-321.
- [22] Lee E. Statistical methods for survival data analysis. A Wiley-Interscience Publication John Wiley & Sons, Inc., USA, 1992.
- [23] Procházka V, Trněný M, Pytlík R, Vasová I, Král Z, Belada D, Kozák T, Kubáčková K, Siffnerová H, Matuška M, Lysý M, Bolomská I, Petráková K, Otavová B, Příbylová J, Švecová J, Papajik T, Hamouzová M, Petrová M, Zapletalová J, Langová K.: Peripheral T-

cell lymphoma, unspecified - the analysis of the data from the Czech Lymphoma Study Group (CLSG) registry. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub. 2007; 151(1):103-7.

[24] Rektorys K. Přehled užití matematiky, SNTL – Nakladatelství technické literatury, Praha, 1981.

[25] Storer B. E., Crowley J.: A diagnostic for Cox regression and general conditional likelihoods. Journal of the American Statistical Association, 1985; 80: 139-147.

[26] Tarone R., Ware J. On distribution free tests for equality of survival distributions. Biometrika, 1977; 64: 156-160.

[27] Tarone R. Tests for trend in life table analysis. Biometrika, 1975; 62: 679-682.

[28] Zheng Z. Two methods of estimating the mean survival time from censored samples. The Indian Journal of Statistics, 1995, A 1: 126-136.

[29] <http://samba.fsv.cuni.cz/~soukup/hendl/csverohodnost.ppt>



## Publikační činnost autora

### Publikace v odborných časopisech s impakt faktorem:

- 1) Šimek M., Němec P., Gwozdziejewicz M., **Langová K.**, Študentová H., Bruk V., Hájek R., Lonský V.: Endoscopic versus minimally invasive vein harvesting. Impact on leg-related morbidity in coronary artery bypass surgery: one-year follow-up of a prospective trial. *J Cardiovasc Surg.* 2008; 49(5):673-678. IF 1,570
- 2) Školoudík D, Bar M, Škoda O., Václavík D., Hradílek P., Allendoerfer J., Šaňák D., Hluštík P., **Langová K.**, Herzig R., Kaňovský P.: Safety and Efficacy of the Sonographic Acceleration of the Middle Cerebral Artery Recanalization: Results of the pilot Thrombotripsy Study. *Ultrasound Med Biol.* 2008; Jun 4 [Epub ahead of print]. IF 2,221
- 3) Šimek M., Němec P., Bruk V., Gwozdziejewicz M., Fluger I., **Langová K.**: Postoperative and Midterm Outcomes of Minimally Invasive and Endoscopic Great Saphenous Vein Harvesting for Coronary Artery Bypass Grafting - A Prospective Analysis. *Thorac Cardiovasc Surg.* 2007; 55(7):428-432 IF 0,814
- 4) Školoudík D., Fadrná T., Bar M., Zapletalová O., Zapletal O., Blatný J., Penka M., **Langová K.**, Hluštík P., Herzig R., Kaňovský P.: Changes in haemocoagulation in healthy volunteers after a 1-hour thrombotripsy using a diagnostic 2-4 MHz transcranial probe. *J Tromb Trombolysis* 2007; (31) [Epub ahead of print]. IF 1,2
- 5) Školoudík D., Bar M., Zapletalová O., **Langová K.**, Herzig R., Kaňovský P.: Hladina D-dimerů u pacientů s akutní ischemickou cévní mozkovou příhodou. *Ces Slov Neurol N* 2007; 70/103(4): 375-379. IF 0,045
- 6) Školoudík D., Fadrná T., Bárťová P., **Langová K.**, Rössner P., Zapletalová O., Hluštík P., Herzig R., Kaňovský P.: Reproducibility of sonographic measurement of the substantia nigra. *Ultrasound Med Biol.* 2007; 33(9):1347-1352. IF 2,221

- 7) Školoudík, M. Bar, O. Škoda, D. Václavík, P. Hradílek, **Langová K.**, Herzig R., Kaňkovský P.: Účinnost sonotrombotripse a sonotrombolýzy v rekanalizaci střední mozkové tepny. Čes. a slov. Neurol. Neurochir. 2007; 70/103(3): 129-133. IF 0,07
- 8) Chlup R., Jelenová D., Kudlová P., Chlupová K., Bartek J., Zapletalová J., **Langová K.**, Chlupová L.: Continuous Glucose Monitoring - A Novel Approach to the Determination of the Glycaemic Index of Foods (DEGIF 1). Exp Clin Endocrinol Diabetes 2006; 114: 68 – 74. IF 1,196
- 9) Školoudík D., Václavík D., Chudoba V., Bar M., **Langová K.**, Herzig R, Kaňovský P.: Riziko vaskulární příhody u pacientů se stenózou intrakraniální tepny. Čes. a slov. Neurol. Neurochir. 2006; 69(102): 355-359. IF 0,07

**Publikace v odborných časopisech a sbornících:**

- 1) **Langová K.**, Zapletalová J., Komenda S.: Statistická analýza přežívání. Sborník příspěvků XXXI. dny lékařské biofyziky, 2008: 86-92; ISBN 978-80-2441985-5
- 2) Komenda S., **Langová K.**: Induktivní statistika – integrující přístup k myšlení člověka bádajícího (Homo scholasticus). Sborník referátů z mezinárodní vědecké konference Současné integrující přístupy k pojetí člověka, 2008: 40-46; ISBN 978-80-86914-33-6
- 3) Galló J., **Langová K.**, Havránek V., Čechová I.: Poor survival of ABG I hip prosthesis in younger patients. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub. 2008; 152(1):163-169.
- 4) Pika T., Minařík J., Schneiderka P., Budiková M., **Langová K.**, Lochman P., Bačovský J., Farbiaková V., Ščudla V.: The correlation of serum immunoglobulin free light chain levels and selected biological markers in multiple myeloma. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub. 2008;152(1): 61-65.

- 5) Köcher M., Černá M., Utíkal P., Kozák J., Šišola I., Bachleda P., Dráč P., Sekanina Z., **Langová K.**: Subintimální angioplastika ve femoropopliteální oblasti – střednědobé výsledky. *Ces Radiol* 2008; 62(2): 153-159.
- 6) Rehák M., Krčová V., Fric E., Slavík L., Bábková B., Prachařová Z., **Langová K.**, Řehák J.: Disturbances of the plasma coagulation defects in retinal venous occlusions. *Cesk Slov Oftalmol.* 2008; 64(3):108-111.
- 7) Rehák M., Fric E., Řehák J., Raiskup-Wolf F., **Langová K.**: Functional examination of retinal vessels in patients with central retinal vein occlusion. *Cesk Slov Oftalmol.* 2007; 63(2):95-102.
- 8) Chlup R., Kudlová P., Peterson K., Slezáková L., Bartek J., Nakládalová M., Sečkař P., Zapletalová J., **Langová K.**, Táncošová S., Doubravová B.: Glykemický index potravin u osob s diabetem 1. a 2. typu a u zdravých. *Diabetes a obezita*, 2007; 13(7): 78-79.
- 9) Procházka V, Trněný M, Pytlik R, Vasová I, Král Z, Belada D, Kozák T, Kubáčková K, Siffnerová H, Matuška M, Lysý M, Bolomská I, Petráková K, Otavová B, Příbylová J, Švecová J, Papajik T, Hamouzová M, Petrová M, Zapletalová J, **Langová K.**: Peripheral T-cell lymphoma, unspecified - the analysis of the data from the Czech Lymphoma Study Group (CLSG) registry. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub.* 2007; 151(1):103-7.
- 10) Nováková K., Polášková B., Fritschová E., **Langová K.**: Úpravy fyziognomie u dětí s morbus Down. *Čes. Stomat.* 2006; 106(4): 115-122.
- 11) Halířová M., Špidlen M., Kamínek M., **Langová K.**: Ztráta kotvení při extrakční ortodontické terapii. *Ortodoncie* 2006;15(5): 30-43.
- 12) Moliková R., Bezdičková M., **Langová K.**, Holibka V., David O., Michalíkova Z., Řehořová J.: The relationship between morphological indicators of human body and posture. *Biomed Pap Med Fac Univ Palacky Olomouc* 2006; 150(2):261-265.

13) Šimek M., Němec P., **Langová K.**: Minimally invasive and endoscopic vein harvesting for coronary artery bypass grafting. *New Frontiers in the Research of PhD Students*, 2006: 169-173, ISBN 80-229-8103-X.

**Publikace abstrakt v časopisech a sbornících:**

- 1) Peterson K., Chlup R., **Langová K.**, Zapletalová J., Kudlová P., Příbylová H.: Benefits of Continuous Glucose Monitoring Transmitted to Paradigm 722, *Diabetes*. 2007; 56(1): A531. IF 7,955
- 2) Chlup R., Řehořová J., **Langová K.**, Zapletalová J., Kudlová P., Příbylová H., Peterson K., Bartek J., Slezáková L.: Effects of Two-Month Consumption of Selected Foods with Low/High Glycemic Index on Metabolic and Anthropometric Parameters in Healthy Persons (DEGIF 5). *Diabetes*. 2007; 56(1):A688. IF 7,955
- 3) Peterson K., Nakládalová M., Chlup R., Kudlová P., Slezáková L., Zapletalová J., **Langová K.**, Doubravová B., Bartek J., Sečkař P.: Influence of Oral Antidiabetic Drugs on Hyperglycemic Power of Foods in Persons with Type 2 Diabetes Mellitus (DEGIF 4). *Diabetes*. 2007; 56(1):A688. IF 7,955
- 4) Vlachová I., Vaverková H., Herzig R., Novotný D., Krčová V., Bártková A., Šaňák D., **Langová K.**, Kaňovský P.: Metabolický syndrom, poruchy fibrolýzy a ultrasenzitivní CRP u nemocných v akutní fázi ICMP. *Česká a slovenská neurologie a neurochirurgie*. 2007; 70(103(Suppl.)): S10. IF 0,045
- 5) Chlup R., Peterson K., Zapletalová J., **Langová K.**, Kudlová P., Příbylová H., Matušková V.: Relation between HbA1c and Average One-Month Interstitial Fluid Glucose Concentration (ISFG) Determined by Continuous Glucose Monitoring - Preliminary Results of an Ongoing Study. *Diabetes*. 2007; 56(1):A527. IF 7,955
- 6) Vlachová I., Vaverková H., Herzig R., Novotný D., Krčová V., Bártková A., Šaňák D., **Langová K.**, Kaňovský P.: Metabolic Syndrome of Insulin Resistance Fibrinolysis and High Sensitivity CRP in the Acute Phase of Ischemic Stroke and their Relation to Carotid Disease. *European Journal of Neurology*. 2007 14 (Suppl. 1), 46. IF 2,437

- 7) Pika T., Minařík J., Zemanová M., Budíková M., **Langová K.**, Bačovský J.: Vztah sérových hladin volných lehkých řetězců imunoglobulinu a vybraných biologických parametrů u mnohočetného myelomu. *Vnitř Lék* 2007; 53(5): 589.
- 8) Matušková J., Hučíková J., Příbylová H., **Langová K.**, Chlup R.: Vztah mezi HbA1c a průměrnou koncentrací P-glukózy v předchozích měsících. *Vnitř Lék* 2007; 53(5): 616.
- 9) Šimek M., Němec P., Bruk V., Marcián P., Hájek R., **Langová K.**: Endoscopic Vein Harvesting in Patients at High Risk of Leg-wound Complications. Annual Meeting of Society for Cardiothoracic Surgery in Great Britain and Ireland, Manchester, March 2007, UK, Book of abstracts, p. 107, No. of abstract 52.
- 10) Pika T., Minařík J., Zemanová M., Budíková M., **Langová K.**, Bačovský J.: Vztah sérových hladin volných lehkých řetězců imunoglobulinu a vybraných biologických parametrů u mnohočetného myelomu. *Vnitř Lék* 2007; 53(5): 589.
- 11) Chlup R., Jelenová D., Chlupová K., Bartek J., Študentová H., Zapletalová J., **Langová K.**, Chlupová L.: Continuous Glucose Monitoring - A New Effective Approach to Determine the Glycemic Index of Foods (DEGIF 1). *Journal of Diabetes Science and Technology* (online). 2007; 1(2): 334.
- 12) Chlup R., Sečkař P., Zapletalová J., **Langová K.**, Kudlová P., Peterson K., Bartek J., Hučíková J.: DegifXL - a new software program for determination of group-related and person-related glycemic indexes of foods by means of continuous glucose monitoring system CGMS. *Acta Diabetologica Romana*. 2007; 33(1):45-46.
- 13) Chlup R., Sečkař P., Kudlová P., Peterson K., **Langová K.**, Bartek J.: Demonstrace nového softwarového programu DegifXL ke stanovení skupinového a individuálního glykemického indexu potravin pomocí CGMS. *Diabetologie Metabolismus Endokrinologie Výživa*. 2007; 10(1):57-58.
- 14) Ščudla V., Pika T., Budíková M., Ordeltová M., Minařík J., Zemanová M., Bačovský J., **Langová K.**: Evaluation of selected biological papers in stages of multiple myeloma

examined according to Durie-Salmon (D-S) and International prognostic index (IPI). *Haematologica: the hematology journal*. 2007; 92(6,supl.2):115-116.

- 15) Chlup R., Peterson K., **Langová K.**, Zapletalová J., Příbylová H., Matušková V., Hučíková J., Kudlová P., Pallayová M.: High Correlation between Hemoglobin Alc and Average 1 - to 3 Month Interstitial Fluid Glucose Concentrations. *Diabetes Technology Society*. 2007; (N/A):A26.
- 16) Vlachová I., Vaverková H., Herzig R., Novotný D., Krčová V., Bártková A., Šaňák D., Kaňovský P., **Langová K.**: Metabolický syndrom, poruchy fibrinolýzy a ultrasenzitivní CRP u nemocných v akutní fázi ischemické cévní mozkové příhody. *Neurologia*. 2007; 2(supl.2): 34.
- 17) Peterson K., Chlup R., **Langová K.**, Zapletalová J., Příbylová H., Hučíková J., Jadviščoková T. Kudlová P; Pallayová M: Reduction of Hemoglobin Alc Resulting from 1 month of Continuous Glucose Monitoring in Persons with Type 1 Diabetes on Paradigm 722. *Diabetes Technology Society*. 2007; A144.
- 18) Nováčková S., Marek I., Kamínek M., **Langová K.**: Tvorba kosti ortodontickým posunem zubu a její stabilita v čase. 8. kongres České ortodontické společnosti, 5.-7.10.2007, Plzeň. 2007:17-18.
- 19) Chlup R., Sečkař P., Zapletalová J., **Langová K.**, Bartek J., Jelenová D.: User-Friendly Software for Determination of glycemic Indexes os Foods by Means of Continuous Glucose Monitoring System (CGMS, Medtronic-Minimed). *Journal of Diabetes Science and Technology (online)*. 2007; 1(2):335.
- 20) Peterson K., **Langová K.**, Kudlová P., Slezáková L., Chlup R: Vliv perorálních antidiabetik na glykemický index potravin u osob s diabetem 2. typu. *Konference vědeckých prací studentů DSP*. 2007:38-39.
- 21) **Langová K.**, Mihál V., Číhalík Č., Potomková J.: A web-based medical education portal: preliminary results. *Bratisl Lek Listy* 2006; 107 (4): 158.

- 22) Chlup R., Zapletalová J., Sečkář P., Malá E., Kudlová P., **Langová K.**: Complementary Therapy with Insulin Aspart vs Regular Human Insulin in Persons with Type 2 Diabetes Mellitus. *Diabetes A Journal of the American Diabetes Association* 2006; 55 suppl 1: A474. IF 8,028
- 23) Pešák J., Grézl T., Zapletalová J., **Langová K.**, Modráčková L.: Decrease of speech non-fluency by bronchodilatation. *Bratisl Lek Listy* 2006; 107 (4): 166 - 167, ISSN 0006-9248
- 24) Chlup R., Jelenová D., Kudlová P., Bartek J., Nakládalová M., Sečkář P., Zapletalová J., **Langová K.**, Pukowietz L., Chlupová K.: Konventionelle Bestimmung des glykämischen Indexes verschiedener Nahrungsmittel (Schokolade, Apfelmuss, Reisbrot, Yoghurt) bei pumpenbehandelten Typ-1 Diabetikern. *Diabetologie* 2006; 1: 421.
- 25) Pika T., Zemanová J., Minařík J., **Langová K.**, Ščudla V.: Vztah sérových hladin volných lehkých řetězců imunoglobulinu ke stupni progresse mnohočetného myelomu. *Vnitř Lék* 2006; 52(11).
- 26) Pírek P., Nováková K., Černochová D., Pospíšilová E., Novotný R, **Langová K.**: Kompozitní korunky v dočasné dentici – prevence progresse kazu. Sympóziium primárnej prevence ústneho zdravia, Zborník príspevkov 2006; S19.

### **Příspěvky na konferencích**

- 1) **Langová K.**, Zapletalová J., Komenda S.: Statistická analýza přežívání. XXXI. dny lékařské biofyziky, 28.-30. května 2008, Malá Morávka.
- 2) Šimek M., Hájek R., Fluger I., **Langová K.**, Lonský V.: Vacuum-assisted closure in the treatment of deep sternal wound infection. 5-years single centre experience. 1st Central European Congress of Surgery, 23-26.4. 2008 Praha.

- 3) Šimek M., Hájek R., Zálešák B., Grulichová J., Němec P., Lonský V., **Langová K.**: Topical negative pressure versus conventional treatment of deep sternal infection in cardiac surgery. EWMA Journal 2008; 8(Suppl):46 18th Conference of the European Wound Management Association, 14-16.5. 2008 Lisabon, Portugalsko.
- 4) Šimek M., Grulichová J., **Langová K.**, Lonský V., Ječmínková L.: Silver-containing hydrofiber dressing versus topical antiseptics in the treatment of surgical site infection in cardiac surgery. Prospective randomized study. EWMA Journal 2008; 8(Suppl):260 18th Conference of the European Wound Management Association, 14-16.5. 2008 Lisabon, Portugalsko.
- 5) Komenda S., **Langová K.**: Induktivní statistika – integrující přístup k myšlení člověka bádajícího. Současné integrující přístupy k pojetí člověka 19. – 20. června 2007 Olomouc.
- 6) Vidura R., Fryšák Z., **Langová K.**: Statisticky prokázané rizikové faktory tranzitorní hypoparatyreózy u nemocných po totálních tyroidektomiích na Klinice ORL a CHHK LF UP a FNO. Olomoucké otolaryngologické odpoledne, 2007, Olomouc.
- 7) Nováčková S., Marek I., Kamínek M., **Langová K.**: Tvorba kostí ortodontickým posunem zubu a její stabilita v čase. 8. kongres České ortodontické společnosti, 5.-7.10.2007 Plzeň.
- 8) Vidura R., Fryšák Z., **Langová K.**: Statisticky prokázané rizikové faktory tranzitorní hypoparatyreózy u nemocných po totálních tyroidektomiích na Klinice ORL a CHHK LF UP v Olomouci a FNO. XXX. Endokrinologické dny, 4.-6. října 2007.
- 9) **Langová K.**, Mihál V., Číhalík Č., Potomková J.: A web-based medical education portal: preliminary results. XXIX. Dni lékařské biofyziky 16. – 18. května 2006 Bratislava.

#### **Odborné přednášky:**

**Langová K.**, Komenda S.: 100 let Studentova t-testu. Přednáška pro Spolek lékařů České lékařské společnosti J. E. Purkyně v Olomouci, 6. 6. 2008 Olomouc.