
Zeit-bis-zum-Ereignis Daten IV

(Vergleich von Überlebenskurven)

Dr. Dirk Hasenclever
IMISE, Universität Leipzig
Dirk.Hasenclever@IMISE.uni-Leipzig.de

Beschreibung von Unterschieden I

Um Unterschiede zwischen Überlebenskurven durch eine eindimensionale Größe zu quantifizieren, betrachten wir einparametrische Familien von Transformationen der y - Achse, bzw. der Zeitachse.

Durch Anwendung der Familie von Transformationen auf eine Basiskurve, entstehen Kurvenscharen, die einen bestimmten Typ von Unterschied darstellen.

(A) Transformationen der y Achse :

$\psi_\theta : [0;1] \rightarrow [0;1]$ monoton $y \mapsto \psi_\theta(y)$ und $\psi_0 = Id_{[0;1]}$

$$S_\theta(t) = \psi_\theta(S_0(t))$$

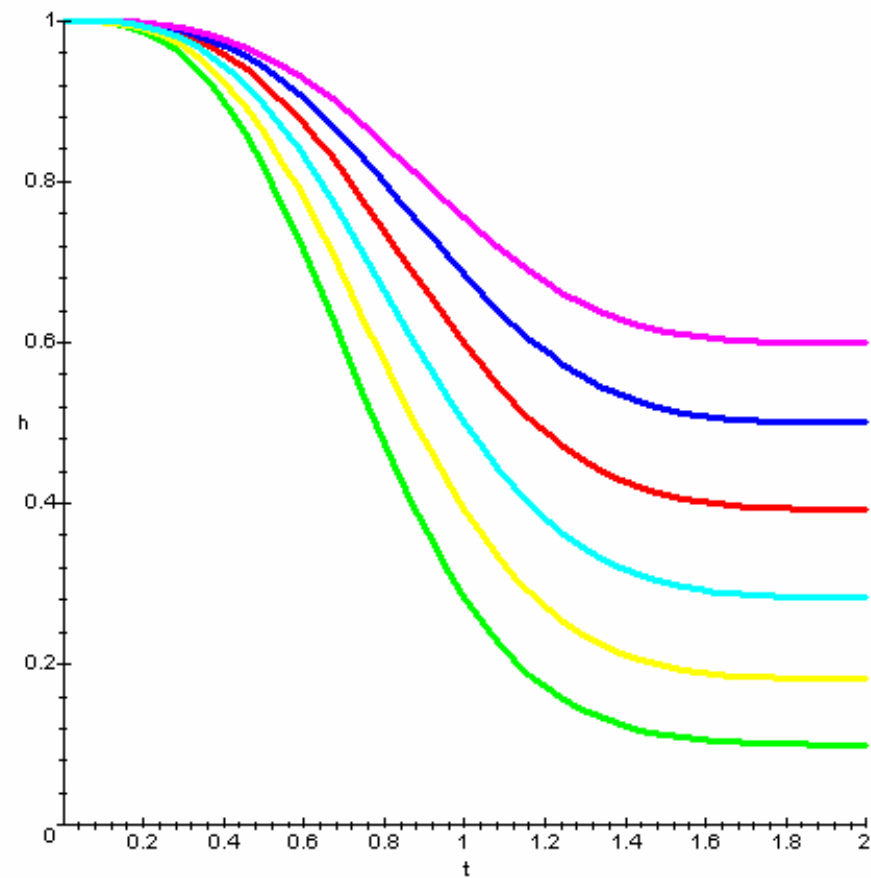
Proportional Hazard

$$S_\theta(t) = S_0(t)^{\exp(\theta)}$$

So genannt wegen $h_\theta(t) = \exp(\theta) \cdot h_0(t)$, $R := \exp(\theta)$ heißt auch "Relatives Risiko"

Proportional Hazard Model:

$$S1(t) := S0(t)^R$$



Beschreibung von Unterschieden II

B Transformationen der t - Achse :

$\psi_\theta : [0; \infty) \rightarrow [0; \infty)$ monoton $y \mapsto \psi_\theta(y)$ und $\psi_0 = Id_{[0;1]}$

$$S_\theta(t) = S_0(\psi_\theta(t))$$

Accelerated failure time model

$$S_\theta(t) = S_0(\theta \cdot t)$$

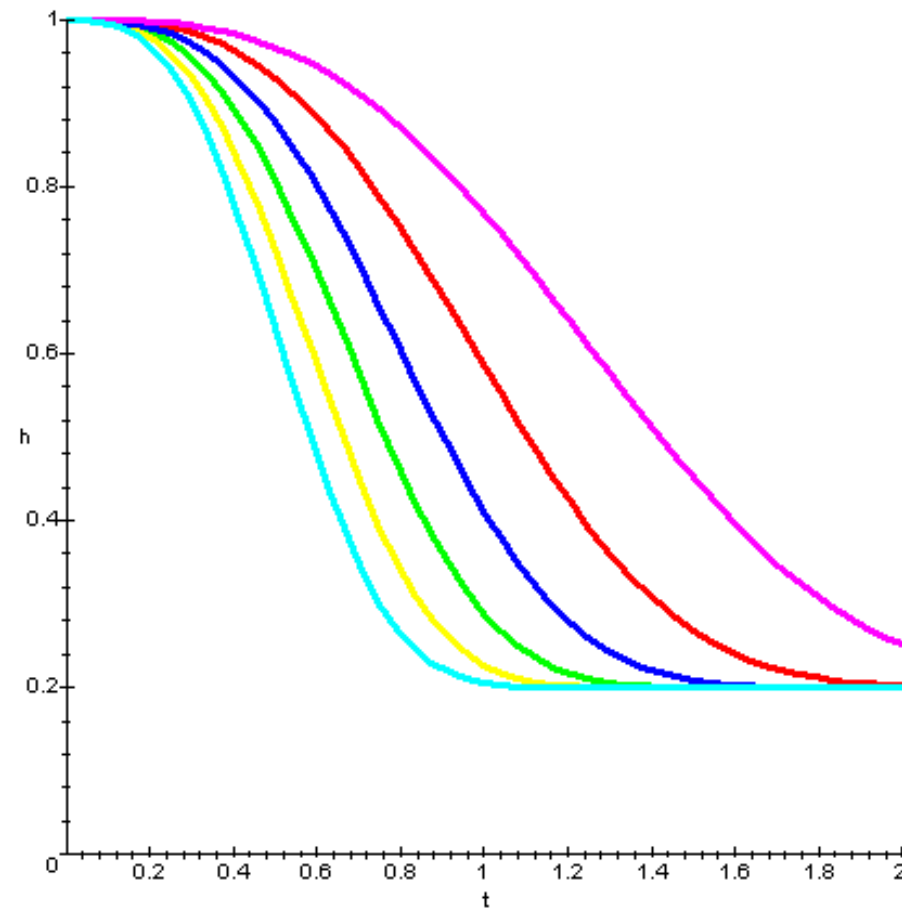
Natürlich können Unterschiede auch durch Überlagerung beider Transformationsarten entstehen.

Proportional Hazard ist häufig approximativ angemessen bei Überlebenskurven in der Onkologie.

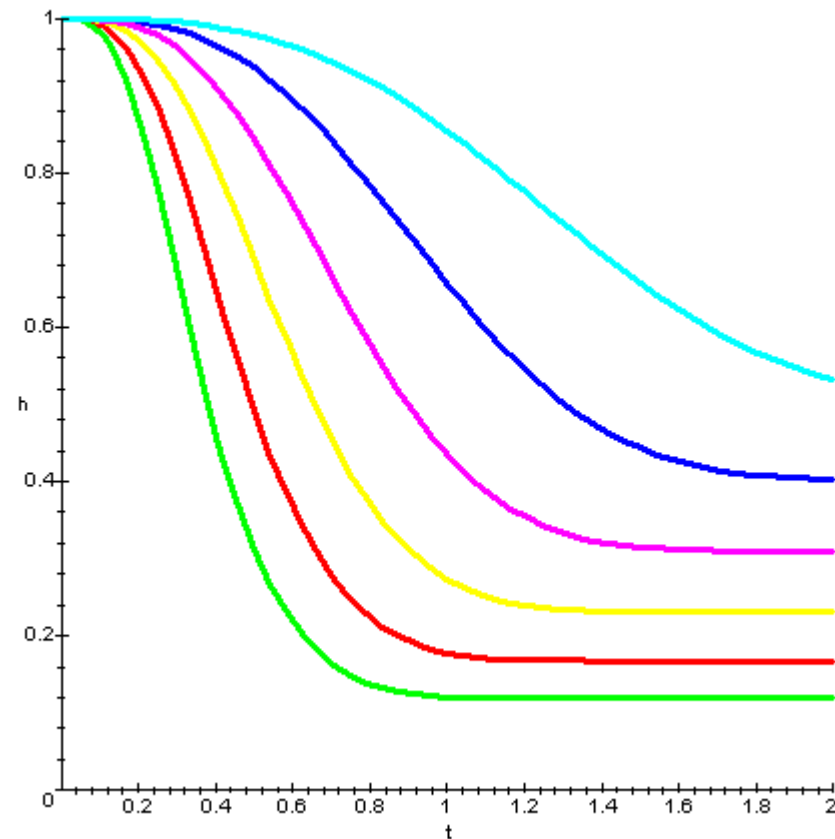
Bei Modulation der Geschwindigkeit eines Prozesses ist das Accelerated failure time model indiziert.

Accelerated failure time model:

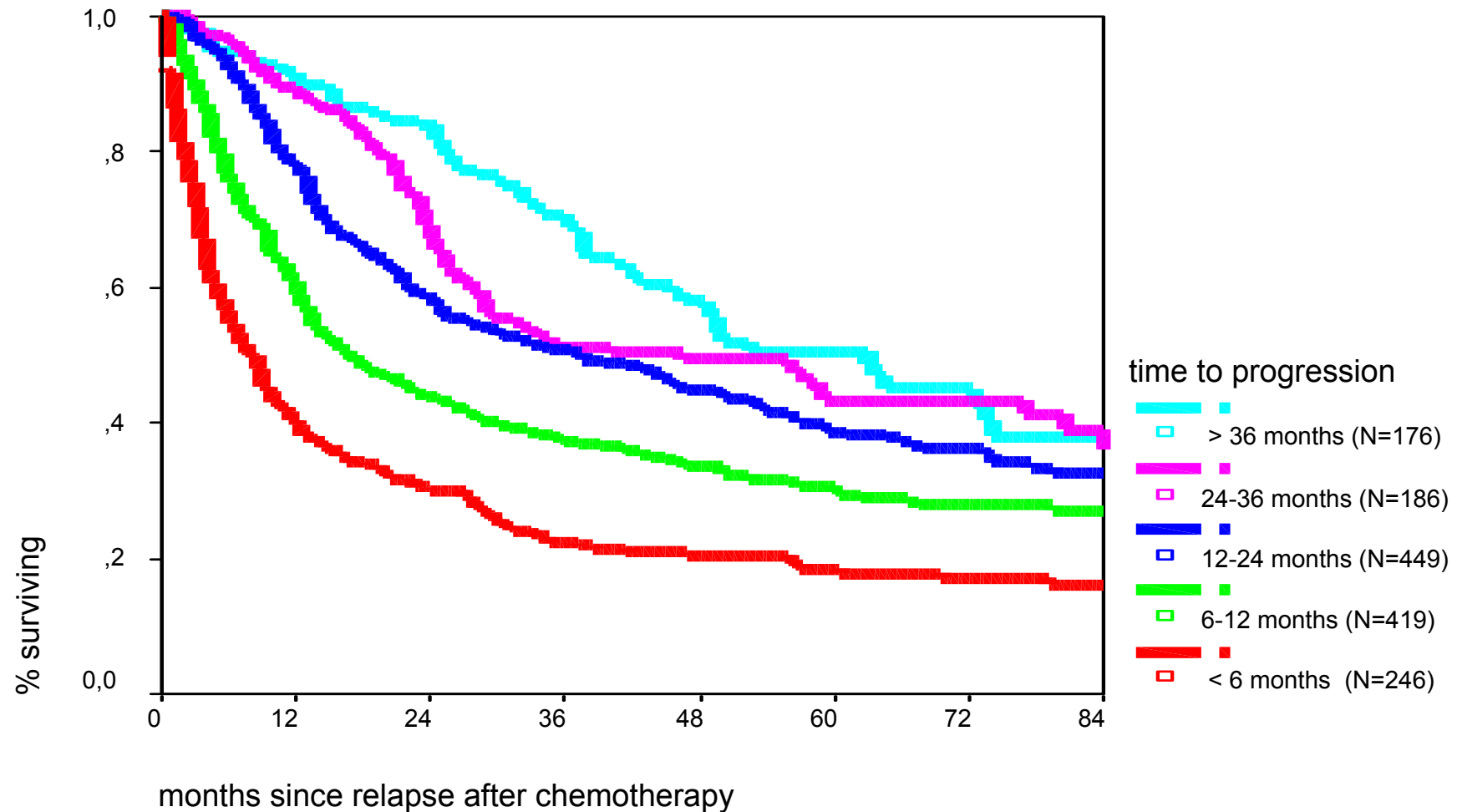
$$S_1(t) := S_0(\gamma * t)$$



Kombination: Proportional Hazard und Accelerated Failure Time



Survival after relapse by time to first chemotherapy failure (N=1476),
Data from the International Prognostic Factor Project on
advanced Hodgkin's disease (IPFP) Hasenclever 2002



Weibull-Verteilung

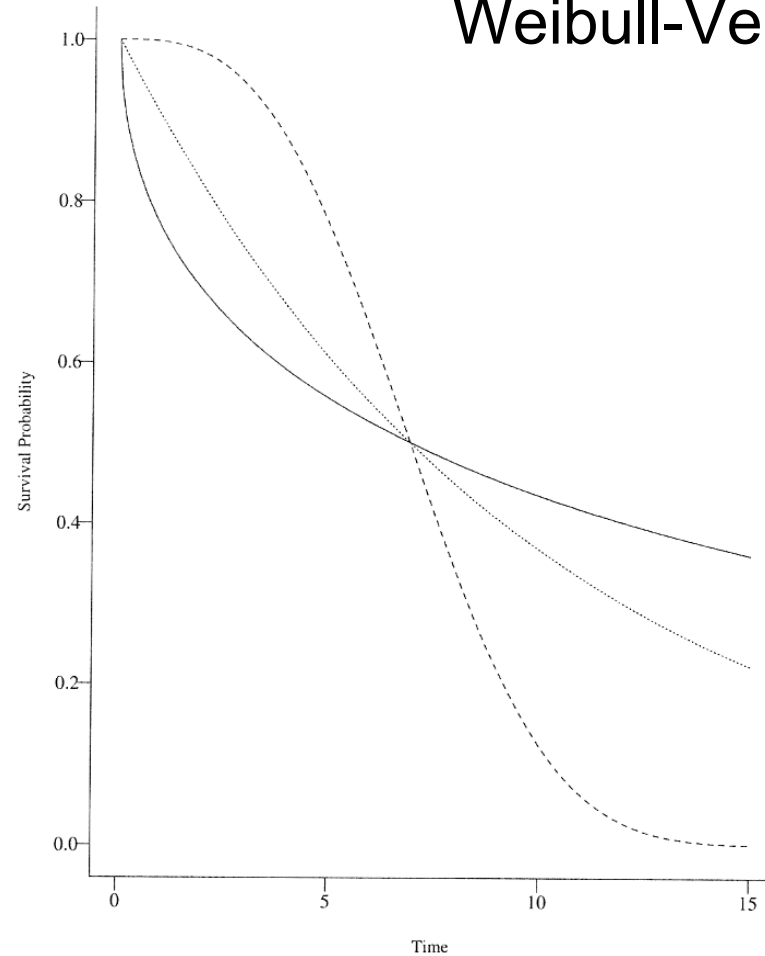


Figure 2.1 Weibull Survival functions for $\alpha = 0.5$, $\lambda = 0.26328$ (—); $\alpha = 1.0$, $\lambda = 0.1$ (·····); $\alpha = 3.0$, $\lambda = 0.00208$ (-----).

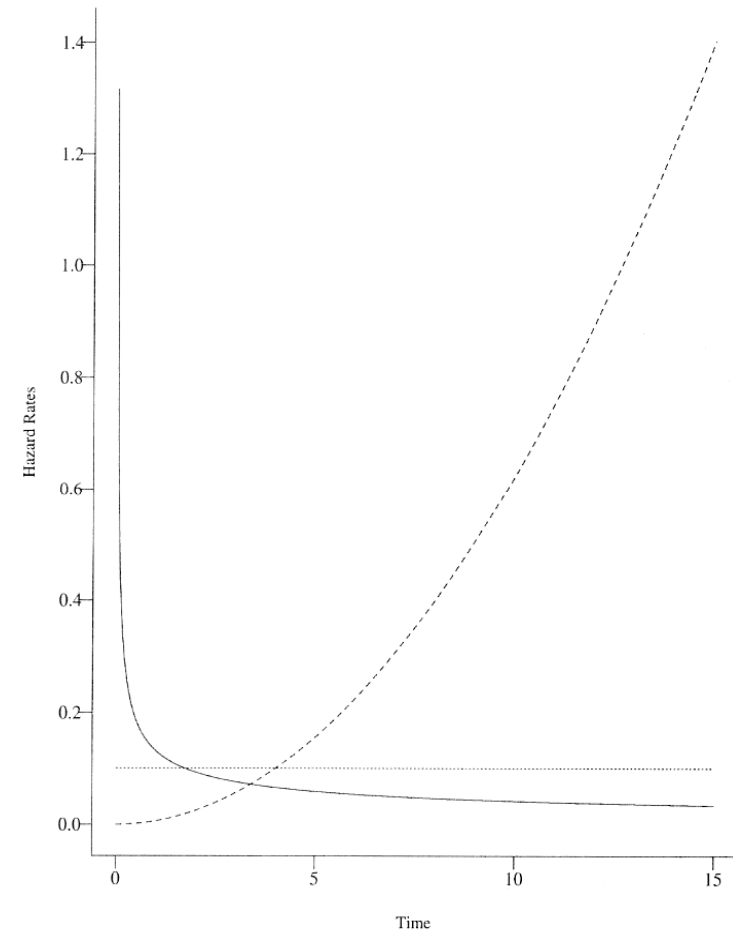


Figure 2.5 Weibull hazard functions for $\alpha = 0.5$, $\lambda = 0.26328$ (—); $\alpha = 1.0$, $\lambda = 0.1$ (-----); $\alpha = 3.0$, $\lambda = 0.00208$ (—·—·—).

$$S(t) = \exp(-\lambda t^\alpha) \text{ für } \lambda, \alpha > 0$$

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

Unterschiede in Weibull-Verteilung

Die WEIBULL - Verteilungsfamilie ist invariant
sowohl unter proportional hazard Transformationen,

$$weibull_{\lambda,\alpha}(t)^R = \exp(-\lambda \cdot t^\alpha)^R = \exp(-(\lambda R) \cdot t^\alpha) = weibull_{(\lambda \cdot R),\alpha}(t)$$

als auch unter accelerated failure time Transformationen

$$weibull_{\lambda,\alpha}(\gamma \cdot t) = \exp(-\lambda \cdot (\gamma \cdot t)^\alpha) = \exp(-(\lambda \cdot \gamma^\alpha) \cdot t^\alpha) = weibull_{(\lambda \cdot \gamma^\alpha),\alpha}(t).$$

Vergleich von Überlebenskurven I

Klasse von Testverfahren zum Vergleich von zwei Überlebenskurven :

Nullhypothese : $H_0 : S_1(t) = S_2(t)$ gegen Alternative : $H_A : S_1(t) \neq S_2(t)$

Seien $t_1 < t_2 < \dots < t_k$ die beobachteten distinkten Ereigniszeitpunkte.

Zum Zeitpunkt t_i werden im j -ten Arm an den n_{ij} Patienten unter Risiko d_{ij} Ereignisse beobachtet.

Sei $d_i = d_{i1} + d_{i2}$ und $n_i = n_{i1} + n_{i2}$.

Unter der Nullhypothese sollte die Wahrscheinlichkeit, dass ein Patient mit Ereignis in t_i zu Arm 1 gehört, proportional zum Anteil n_{i1}/n_i sein.

Vergleich von Überlebenskurven II

Die Teststatistik :

$$Z_1 := \sum_{i=1}^k w_i (\text{Daten}[t_j < t_i]) \left(d_{i1} - \frac{n_{i1}}{n_i} d_i \right)$$

misst daher die Abweichung von der Erwartung unter der Nullhypothese.

w_i ist ein geeignet zu wählendes Gewicht, welches eventuell von den Daten abhängen kann.

Vergleich von Überlebenskurven III

Unter der Nullhypothese hat die Teststatistik Z_1 den Erwartungswert 0.

Konditional auf n_{i1} , n_i und d_i ist d_{i1} unter der Nullhypothese verteilt nach

$$d_{i1} \sim \text{binomial}(d_i, \frac{n_{i1}}{n_i}) \quad \text{mit} \quad \text{var}(d_{i1}) = \frac{n_{i1}}{n_i} \cdot \left(1 - \frac{n_{i1}}{n_i}\right)$$

Als Linearkombination von binomial - verteilten Größen ist Z asymptotisch normalverteilt mit approximativer Varianz :

$$\begin{aligned} \text{var}(Z_1) &\approx \text{var}\left(\sum_i w_i \left(d_{i1} - \frac{n_{i1}}{n_i} d_i\right)\right) = \sum_i w_i^2 \text{var}\left(d_{i1} - \frac{n_{i1}}{n_i} d_i\right) = \sum_i w_i^2 \text{var}(d_{i1}) \\ &= \sum_i w_i^2 \frac{n_{i1}}{n_i} \cdot \left(1 - \frac{n_{i1}}{n_i}\right) \end{aligned}$$

Die Approximation kann man zur Berücksichtigung von sogenannten "ties" ($d_i > 1$)

noch verbessern

$$\text{var}(Z_1) \approx \sum_i w_i^2 \frac{n_{i1}}{n_i} \cdot \left(1 - \frac{n_{i1}}{n_i}\right) \left[\frac{n_i - d_i}{n_i - 1} \right]$$

Vergleich von Überlebenskurven IV

Die Teststatistik

$$Z := \frac{Z_1}{\sqrt{\text{var}(Z_1)}} \sim \text{Normal}(0,1)$$

ist also unter der Nullhypothese
approximativ standardnormalverteilt.

Ein Test z.B. zum 5% Niveau erhält man durch die Regel :
Lehne H_0 ab, falls $|Z| > 1.96$

Vergleich von Überlebenskurven V

Die Wahl der Gewichte w_i erlaubt es einem, Gewicht mehr auf die frühen oder späten Beobachtungen zu legen und so je nach erwartetem Typ von Unterschied die Power zu optimieren.

Der am meisten verwendete Logrank Test verwendet $w_i = 1$.

Die Gewichte (Gewichtsfunktion) sollten vor Inspektion der Daten festgelegt werden.

Warum zeigt das folgende Beispiel.

Zeit bis zur Katheter-Infektion bei Nieren-Dialyse Patienten

In a study (Nahman et al., 1992) designed to assess the time to first exit-site infection (in months) in patients with renal insufficiency, 43 patients utilized a surgically placed catheter (Group 1), and 76 patients utilized a percutaneous placement of their catheter (Group 2). Catheter failure was the primary reason for censoring. Cutaneous exit-site infection was defined as a painful cutaneous exit site and positive cultures, or peritonitis, defined as a presence of clinical symptoms, elevated peritoneal dialytic fluid, elevated white blood cell count (100 white blood cells/ μ l with >50% neutrophils), and positive peritoneal dialytic fluid cultures. The data appears in Table 1.2.

TABLE 1.2

Times to infection (in months) of kidney dialysis patients with different catheterization procedures

<i>Surgically Placed Catheter</i>
<i>Infection Times:</i> 1.5, 3.5, 4.5, 4.5, 5.5, 8.5, 8.5, 9.5, 10.5, 11.5, 15.5, 16.5, 18.5, 23.5, 26.5
<i>Censored Observations:</i> 2.5, 2.5, 3.5, 3.5, 3.5, 4.5, 5.5, 6.5, 6.5, 7.5, 7.5, 7.5, 7.5, 8.5, 9.5, 10.5, 11.5, 12.5, 12.5, 13.5, 14.5, 14.5, 21.5, 21.5, 22.5, 22.5, 25.5, 27.5
<i>Percutaneous Placed Catheter</i>
<i>Infection Times:</i> 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 2.5, 2.5, 3.5, 6.5, 15.5
<i>Censored Observations:</i> 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 1.5, 1.5, 1.5, 1.5, 2.5, 2.5, 2.5, 2.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 4.5, 4.5, 4.5, 5.5, 5.5, 5.5, 5.5, 5.5, 5.5, 6.5, 7.5, 7.5, 7.5, 8.5, 8.5, 8.5, 9.5, 9.5, 10.5, 10.5, 10.5, 11.5, 11.5, 12.5, 12.5, 12.5, 12.5, 14.5, 14.5, 16.5, 16.5, 18.5, 19.5, 19.5, 19.5, 20.5, 22.5, 24.5, 25.5, 26.5, 26.5, 28.5

Zeit bis zur Katheter-Infektion bei Nieren-Dialyse Patienten

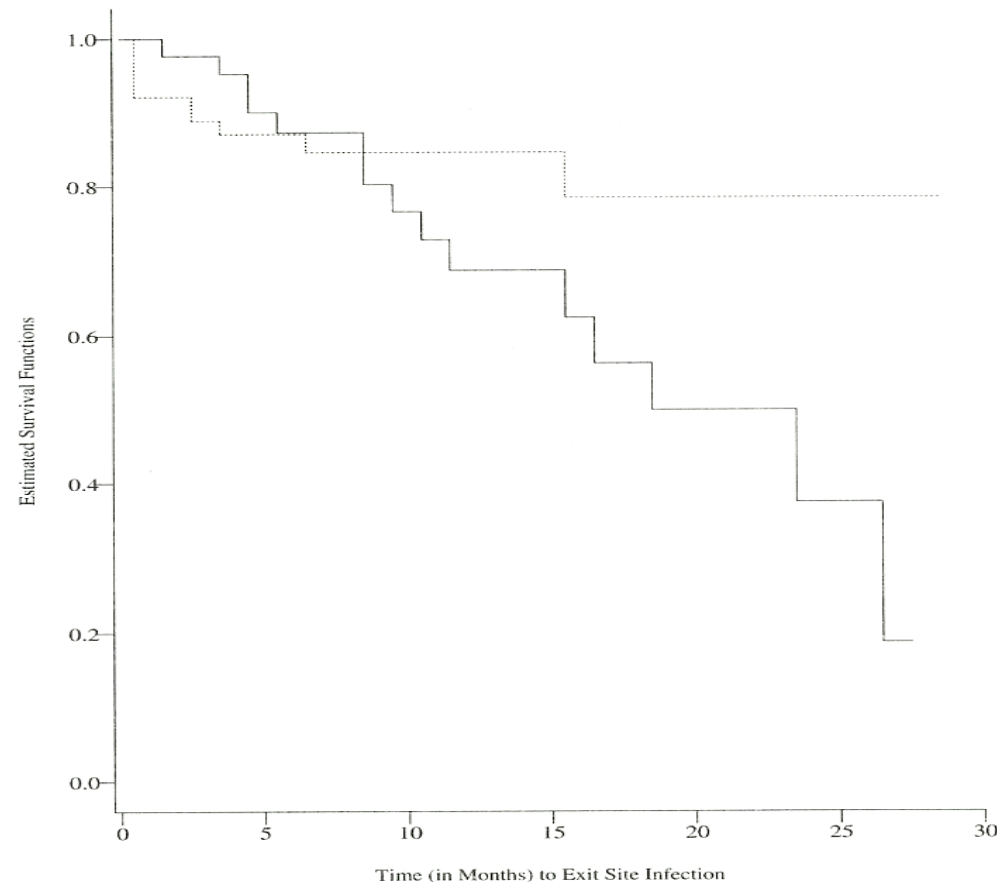


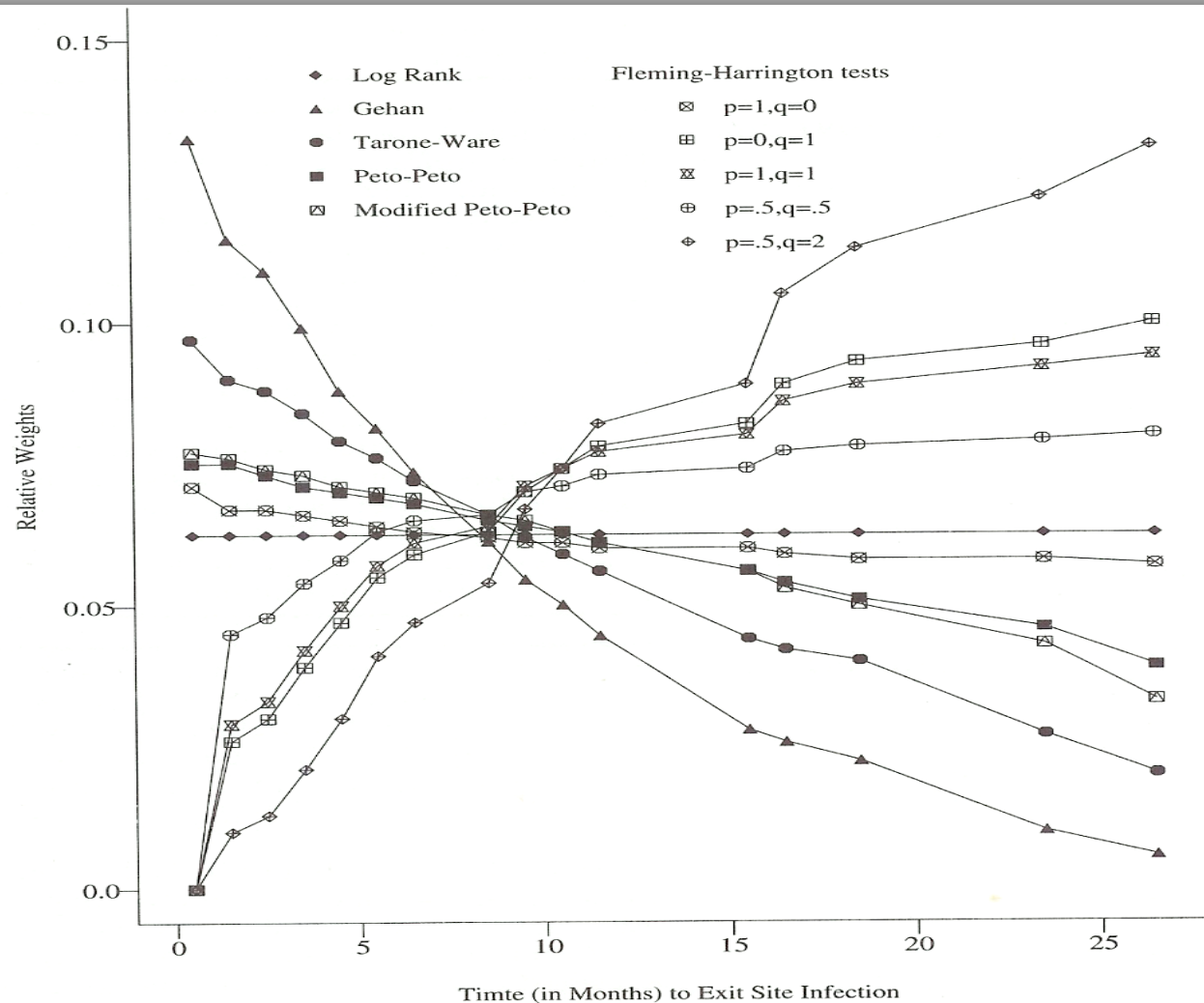
Figure 7.1 *Estimated (Infection-free) survival function for kidney dialysis patients with percutaneous (-----) and surgical (——) placements of catheters.*

Berechnung der Logrank-Statistik

t_i	Y_{i1}	d_{i1}	Y_{i2}	d_{i2}	Y_i	d_i	$Y_{i1} \left(\frac{d_i}{Y_i} \right)$	$d_{i1} - Y_{i1} \left(\frac{d_i}{Y_i} \right)$	$\frac{Y_{i1}}{Y_i} \left(1 - \frac{Y_{i1}}{Y_i} \right) \left(\frac{Y_i - d_i}{Y_i - 1} \right) d_i$
0.5	43	0	76	6	119	6	2.168	-2.168	1.326
1.5	43	1	60	0	103	1	0.417	0.583	0.243
2.5	42	0	56	2	98	2	0.857	-0.857	0.485
3.5	40	1	49	1	89	2	0.899	0.101	0.489
4.5	36	2	43	0	79	2	0.911	1.089	0.490
5.5	33	1	40	0	73	1	0.452	0.548	0.248
6.5	31	0	35	1	66	1	0.470	-0.470	0.249
8.5	25	2	30	0	55	2	0.909	1.091	0.487
9.5	22	1	27	0	49	1	0.449	0.551	0.247
10.5	20	1	25	0	45	1	0.444	0.556	0.247
11.5	18	1	22	0	40	1	0.450	0.550	0.248
15.5	11	1	14	1	25	2	0.880	0.120	0.472
16.5	10	1	13	0	23	1	0.435	0.565	0.246
18.5	9	1	11	0	20	1	0.450	0.550	0.248
23.5	4	1	5	0	9	1	0.444	0.556	0.247
26.5	2	1	3	0	5	1	0.400	0.600	0.240
SUM		15		11		26	11.036	3.964	6.211

Gewichtete Log-rank Statistiken

Relative Gewichte im Katheterbsp.

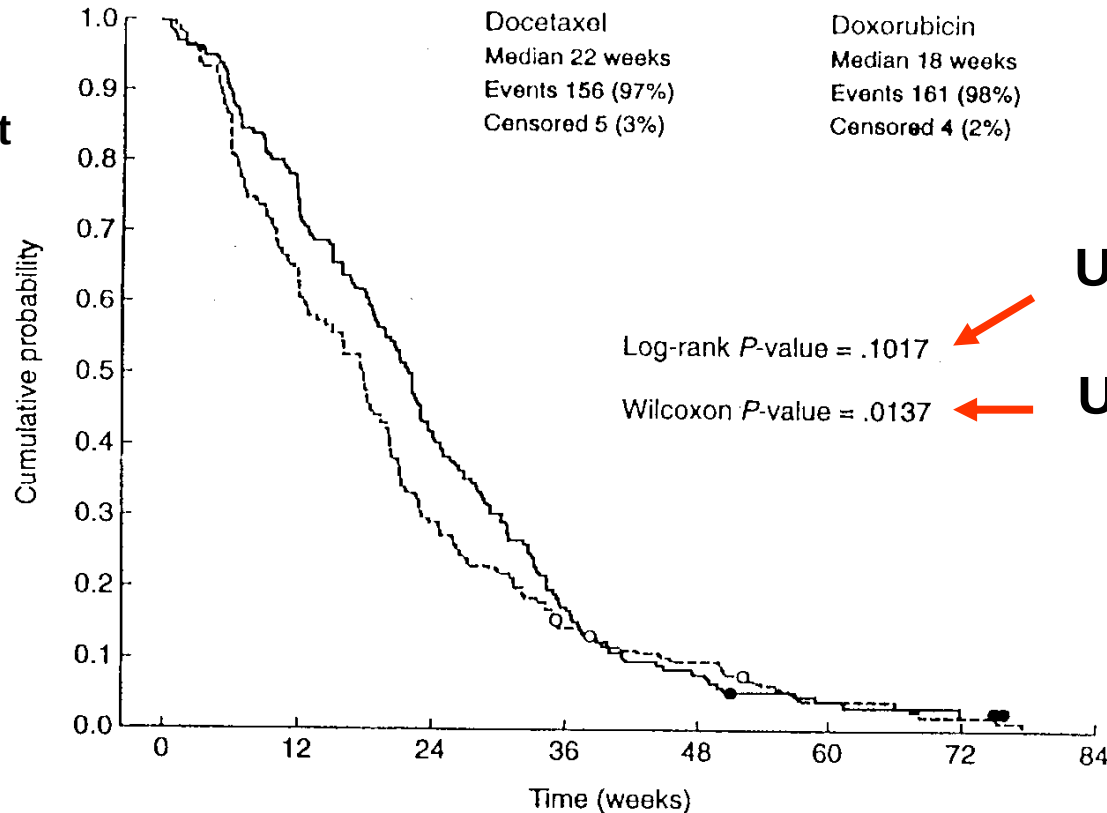


Vergleich der gewichteten Log-Rank Statistiken (Katheterbsp.)

<i>Test</i>	$W(t_i)$	$Z_1(\tau)$	σ_{11}^2	χ^2	<i>p-value</i>
Log-Rank	1.0	3.96	6.21	2.53	0.112
Gehan	Y_i	-9	38862	0.002	0.964
Tarone-Ware	$Y_i^{1/2}$	13.20	432.83	0.40	0.526
Peto-Peto	$\hat{S}(t_i)$	2.47	4.36	1.40	0.237
Modified Peto-Peto	$\hat{S}(t_i) Y_i / (Y_i + 1)$	2.31	4.20	1.28	0.259
Fleming-Harrington $p = 0, q = 1$	$[1 - \hat{S}(t_{i-1})]$	1.41	0.21	9.67	0.002
Fleming-Harrington $p = 1, q = 0$	$\hat{S}(t_{i-1})$	2.55	4.69	1.39	0.239
Fleming-Harrington $p = 1, q = 1$	$\hat{S}(t_{i-1})[1 - \hat{S}(t_{i-1})]$	1.02	0.11	9.83	0.002
Fleming-Harrington $p = 0.5, q = 0.5$	$\hat{S}(t_{i-1})^{0.5}[1 - \hat{S}(t_{i-1})]^{0.5}$	2.47	0.66	9.28	0.002
Fleming-Harrington $p = 0.5, q = 2$	$\hat{S}(t_{i-1})^{0.5}[1 - \hat{S}(t_{i-1})]^2$	0.32	0.01	8.18	0.004

Vorsicht: Testergebnis hängt von der Wahl der Teststatistik (=Differenzmaß) ab

Freedom
From
treatment
failure



Number of patients

Docetaxel	161	123	68	28	13	6	4
Doxorubicin	165	106	48	23	15	6	3

Chan et al: JCO 17,8: 2341-54