
Zeit-bis-zum-Ereignis Daten (Überlebenszeitdaten) I

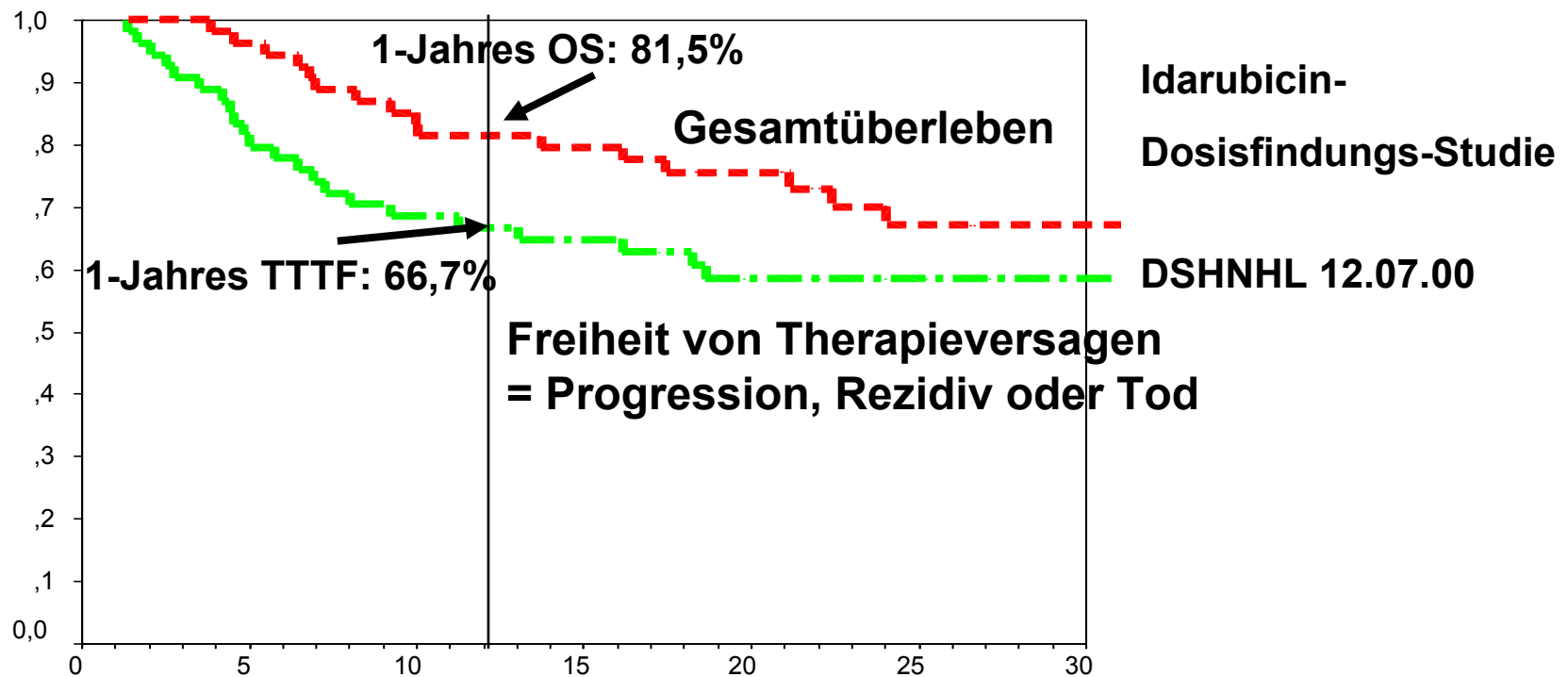
Dr. Dirk Hasenclever

„Zeit bis zum Ereignis“ – Daten I

- Zeit von einem definierten Beginn bis zum Eintreten eines definierten Ereignisses
 - Diagnose bis Tod
 - Diagnose bis Rezidiv
 - Knochenmarktransplantation bis Erholung des Blutbilds
 - OP bis Ausfall des Implantats/Zahns etc.
 - ...
- Formal also eine positive metrische Zufallsgröße $T > 0$

„Zeit bis Ereignis“ – Daten II

- Beschreibung durch invers-kumulierte Verteilungskurven = „Überlebenskurven“



„Zeit bis Ereignis“ – Daten III

- Problem: zensierte Daten
 - Nicht bei allen Patienten wird das Ereignis beobachtet. (z.B.: geheilte Patienten!)
 - Gegeben nur:
 - Zeit bis **Ereignis oder bis zur letzten Beobachtung**
 - Status-Indikator: **Ereignis eingetreten ja/nein**

Darstellung durch (invers kumulierte) Verteilungsfunktion = „Überlebenskurve“

Zeit

pancr598.sav - SPSS Daten-Editor

Datei Bearbeiten Ansicht Daten Transformieren

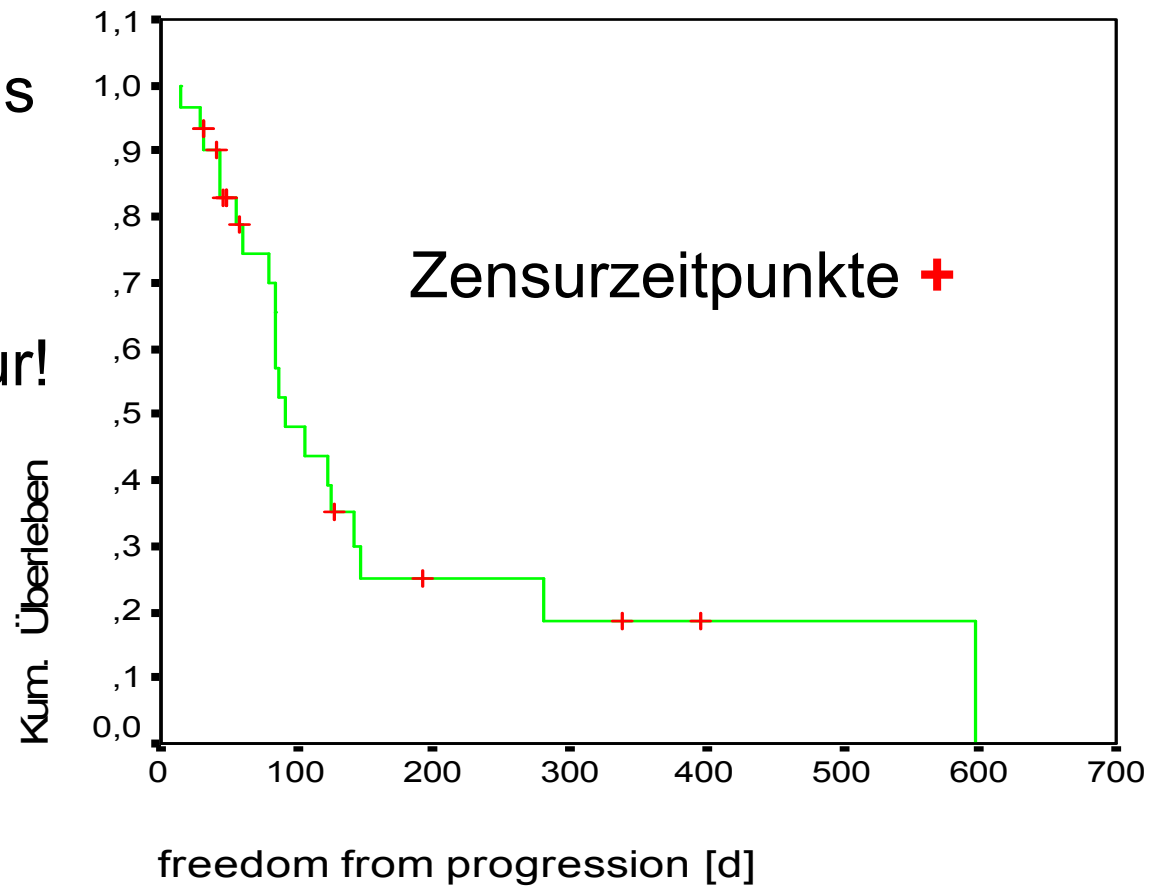
1: ffp 60

	ffp	status	var
1	60,0	1,0000	
2	597,0	1,0000	
3	49,0	,0000	
4	128,0	,0000	
5	338,0	,0000	
6	49,0	,0000	
7	125,0	1,0000	
8	191,0	,0000	
9	280,0	1,0000	
10	90,0	1,0000	
11	146,0	1,0000	
12	40,0	,0000	
13	58,0	,0000	
14	122,0	1,0000	
15	14,0	1,0000	
16	86,0	1,0000	
17	396,0	,0000	
18	78,0	1,0000	
19	142,0	1,0000	
20	29,0	1,0000	
21	106,0	1,0000	
22	31,0	,0000	
23	42,0	1,0000	
24	31,0	,0000	
25	56,0	1,0000	
26	42,0	1,0000	
27	84,0	1,0000	
28	45,0	,0000	
29	84,0	1,0000	
30	32,0	1,0000	
31	83,0	1,0000	
32			

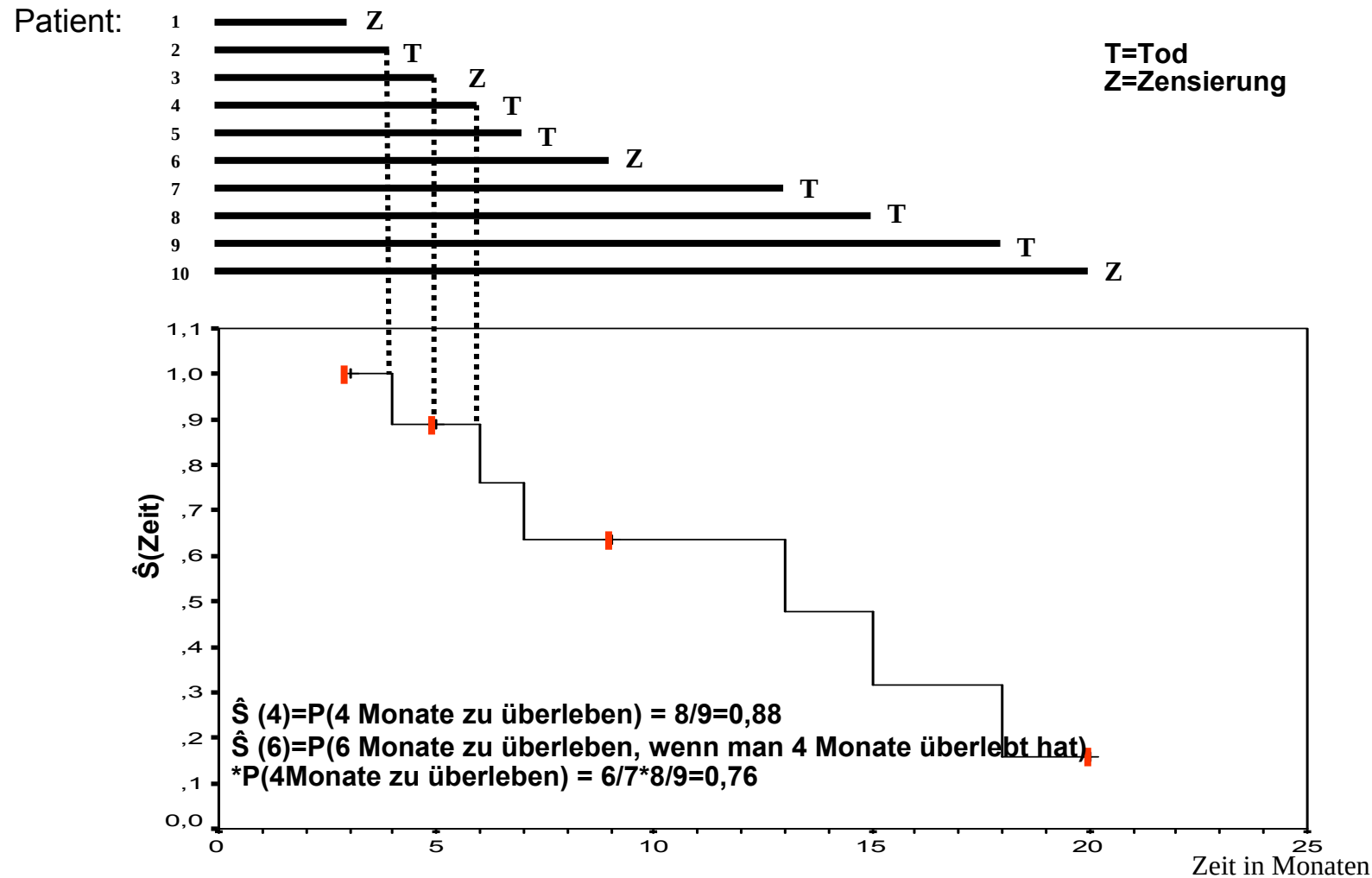
Status

Daten-
Struktur!

Patienten mit Pankreas Ca



Kaplan-Meier-Schätzer $\hat{S}(t)$ für $S(t)$



Übersicht: Zensur-Typen

- $T = t$ (reguläre Beobachtung)
- $T \in (C, \infty]$ (rechtszensiert, C = Zeit letzter Beobachtung)
- $T \in [0, C)$ (linkszensiert, C = Zeit erster Beobachtung)
- $T \in (C_1, C_2)$ (intervallzensiert)

Im Weiteren behandelt werden nur Techniken zum Umgang mit rechts-zensierten Daten als häufigstem Zensurtyp.

Mathematische Beschreibung von Überlebenszeiten: Definitionen

$S(t)$ = invers kumulierte
Verteilungsfunktion

$F(t)$ = kumulierte
Verteilungsfunktion

$f(t) dt$ = Verteilungsdichte

$h(t)$ = Hasardfunktion
(instantane Ereignisrate)

$H(t)$ = kumulative
Hasardfunktion

$$S(t) = 1 - F(t) = \int_t^{\infty} f(\tau) d\tau$$

$$f(t) = F(t)' = -S(t)'$$

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} \ln(S(t))$$

$$H(t) = \int_0^t h(\tau) d\tau = -\ln(S(t))$$

Hasardfunktion I

Bedingte Ereignisrate : $\frac{\text{pr}(t \leq T \leq t + \Delta t \mid T \geq t)}{\Delta t}$

z.B. Sterberate zwischen 45. und 46. Geburtstag ($\Delta t = 1$ Jahr)

Mit $\Delta t \rightarrow 0$ erhält man instantane bedingte Ereignisrate = Hazardfunktion :

$$h(t) = \lim_{\Delta t \rightarrow 0} \left(\frac{\text{pr}(t \leq T \leq t + \Delta t \mid T \geq t)}{\Delta t} \right)$$

Hasardfunktion II

Falls $S(t) = \int_t^{\infty} f(\tau) d\tau$ so ist

$$\begin{aligned} & \lim_{\Delta t \rightarrow 0} \frac{pr(t \leq T \leq t + \Delta t \mid T \geq t)}{\Delta t} \\ &= \frac{\lim_{\Delta t \rightarrow 0} \left(\frac{S(t) - S(t + \Delta t)}{\Delta t} \right)}{S(t)} = -\frac{S(t)'}{S(t)} = \frac{f(t)}{S(t)} \end{aligned}$$

Also

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} [\ln(S(t))]$$

Hasardfunktion III

Zusammenhang zwischen Hasardfunktion und Überlebensfunktion :

$$S(t) = \exp\left(-\int_0^t h(\tau) d\tau\right)$$

$h(t)$ ist die "infinitesimale Wahrscheinlichkeit" im Zeitraum t bis $t + dt$ zu sterben.

$S(t)$ kann man als Produktintegral verstehen,
nämlich als Wahrscheinlichkeit alle infinitesimalen Zeitintervalle dt zu überleben,
gegeben, dass man jeweils deren Anfang erlebt :

$$S(t) = \prod_0^t (1 - h(\tau) d\tau) = \exp\left(\ln\left(\prod_0^t (1 - h(\tau) d\tau)\right)\right) = \exp\left(\int_0^t \ln(1 - h(\tau) d\tau)\right) = \exp\left(-\int_0^t h(\tau) d\tau\right)$$

Produktintegral $\ln(1+x) \approx x$

TABLE 2.1*Survival Functions of U.S. Population By Race and Sex in 1989*

Age	White Male	White Female	Black Male	Black Female	Age	White Male	White Female	Black Male	Black Female
0	1.00000	1.00000	1.00000	1.00000	43	0.93771	0.97016	0.85917	0.93361
1	0.99092	0.99285	0.97996	0.98283	44	0.93477	0.96862	0.85163	0.92998
2	0.99024	0.99232	0.97881	0.98193	45	0.93161	0.96694	0.84377	0.92611
3	0.98975	0.99192	0.97792	0.98119	46	0.92820	0.96511	0.83559	0.92201
4	0.98937	0.99160	0.97722	0.98059	47	0.92450	0.96311	0.82707	0.91761
5	0.98905	0.99134	0.97664	0.98011	48	0.92050	0.96091	0.81814	0.91301
6	0.98877	0.99111	0.97615	0.97972	49	0.91617	0.95847	0.80871	0.90801
7	0.98850	0.99091	0.97571	0.97941	50	0.91148	0.95575	0.79870	0.90271
8	0.98825	0.99073	0.97532	0.97915	51	0.90639	0.95273	0.78808	0.89701
9	0.98802	0.99056	0.97499	0.97892	52	0.90086	0.94938	0.77685	0.89101
10	0.98782	0.99041	0.97472	0.97870	53	0.89480	0.94568	0.76503	0.88451
11	0.98765	0.99028	0.97449	0.97847	54	0.88810	0.94161	0.75268	0.87751
12	0.98748	0.99015	0.97425	0.97823	55	0.88068	0.93713	0.73983	0.87001
13	0.98724	0.98999	0.97392	0.97796	56	0.87250	0.93222	0.72649	0.86191
14	0.98686	0.98977	0.97339	0.97767	57	0.86352	0.92684	0.71262	0.85331
15	0.98628	0.98948	0.97258	0.97735	58	0.85370	0.92096	0.69817	0.84331
16	0.98547	0.98909	0.97145	0.97699	59	0.84299	0.91455	0.68308	0.83331
17	0.98445	0.98862	0.97002	0.97658	60	0.83135	0.90756	0.66730	0.82231
18	0.98326	0.98809	0.96829	0.97612	61	0.81873	0.89995	0.65083	0.81031
19	0.98197	0.98755	0.96628	0.97559	62	0.80511	0.89169	0.63368	0.79731
20	0.98063	0.98703	0.96403	0.97498	63	0.79052	0.88275	0.61584	0.78331
21	0.97924	0.98654	0.96151	0.97429	64	0.77501	0.87312	0.59732	0.76831
22	0.97780	0.98607	0.95873	0.97352	65	0.75860	0.86278	0.57813	0.75231
23	0.97633	0.98561	0.95575	0.97267	66	0.74131	0.85169	0.55829	0.73731
24	0.97483	0.98514	0.95267	0.97174	67	0.72309	0.83980	0.53783	0.72131
25	0.97332	0.98466	0.94954	0.97074	68	0.70383	0.82702	0.51679	0.70331
26	0.97181	0.98416	0.94639	0.96967	69	0.68339	0.81324	0.49520	0.68631
27	0.97029	0.98365	0.94319	0.96852	70	0.66166	0.79839	0.47312	0.66731
28	0.96876	0.98312	0.93989	0.96728	71	0.63865	0.78420	0.45058	0.64731
29	0.96719	0.98257	0.93642	0.96594	72	0.61441	0.76522	0.42765	0.62731
30	0.96557	0.98199	0.93273	0.96448	73	0.58897	0.74682	0.40442	0.60731
31	0.96390	0.98138	0.92881	0.96289	74	0.56238	0.72716	0.38100	0.58731
32	0.96217	0.98073	0.92466	0.96118	75	0.53470	0.70619	0.35749	0.56731
33	0.96038	0.98005	0.92024	0.95934	76	0.50601	0.68387	0.33397	0.53731
34	0.95852	0.97933	0.91551	0.95740	77	0.47641	0.66014	0.31050	0.51731
35	0.95659	0.97858	0.91044	0.95536	78	0.44604	0.63494	0.28713	0.48731
36	0.95457	0.97779	0.90501	0.95321	79	0.41503	0.60822	0.26391	0.46731
37	0.95245	0.97696	0.89922	0.95095	80	0.38355	0.57991	0.24091	0.43731
38	0.95024	0.97607	0.89312	0.94855	81	0.35178	0.54997	0.21819	0.40731
39	0.94794	0.97510	0.88677	0.94598	82	0.31991	0.51835	0.19583	0.37731
40	0.94555	0.97404	0.88021	0.94321	83	0.28816	0.48502	0.17392	0.34731
41	0.94307	0.97287	0.87344	0.94023	84	0.25677	0.44993	0.15257	0.31731
42	0.94047	0.97158	0.86643	0.93703	85	0.22599	0.41306	0.13191	0.28731

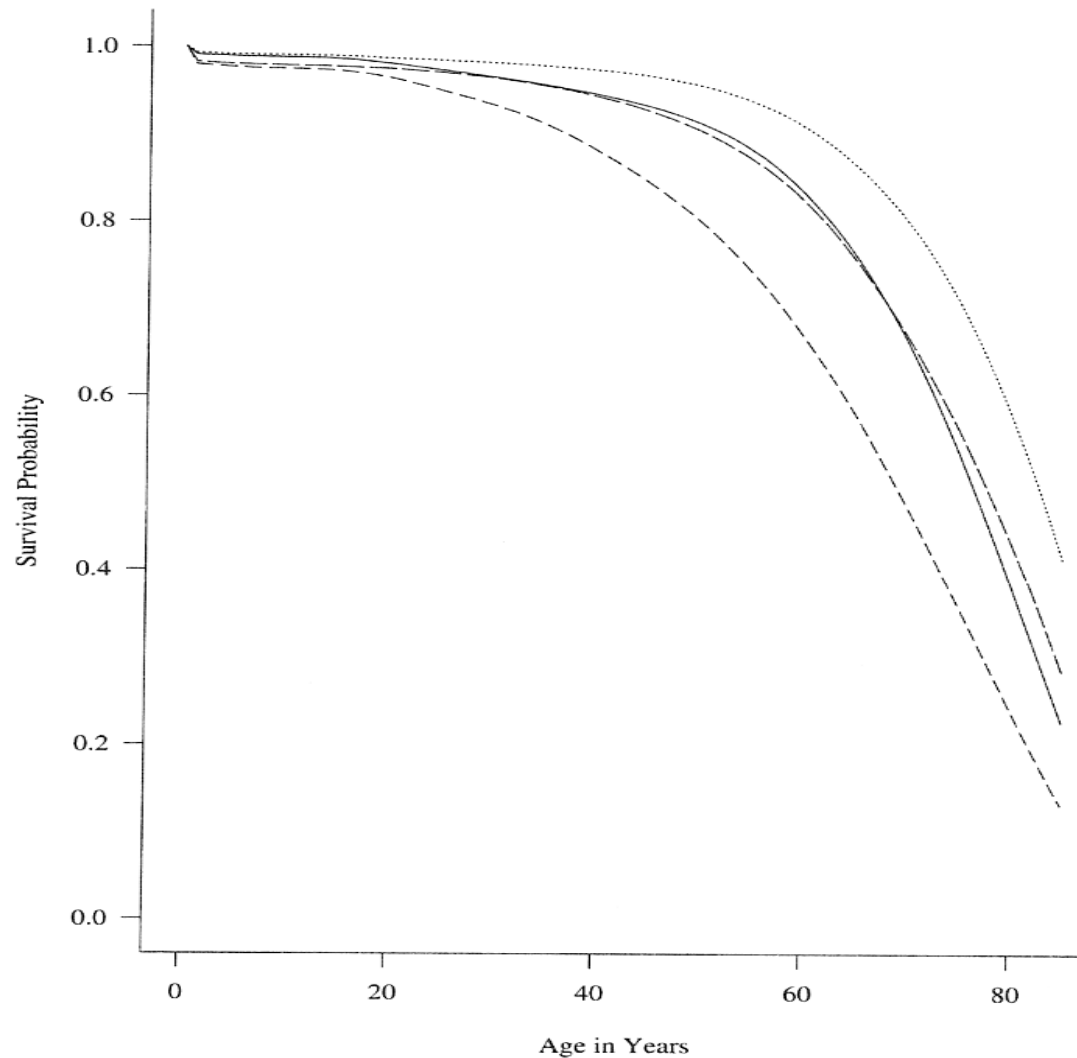


Figure 2.2 *Survival Functions for all cause mortality for the US population in 1989. White males (—); white females (·····); black males (-----); black females (— — — —).*

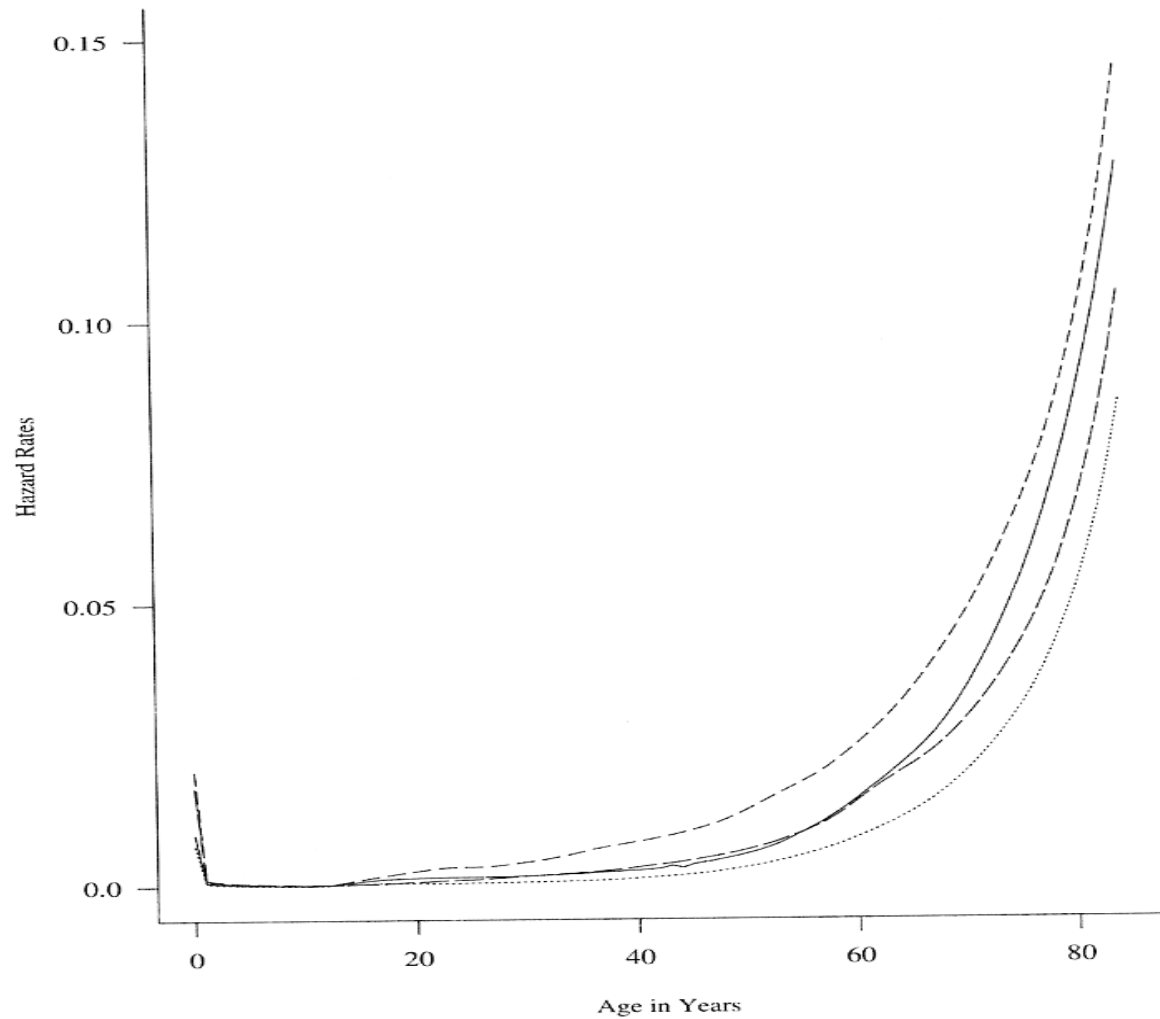


Figure 2.6 *Hazard functions for all cause mortality for the US population 1989. White males (—); white females (·····); black males (- - - - - black females (— — —)).*

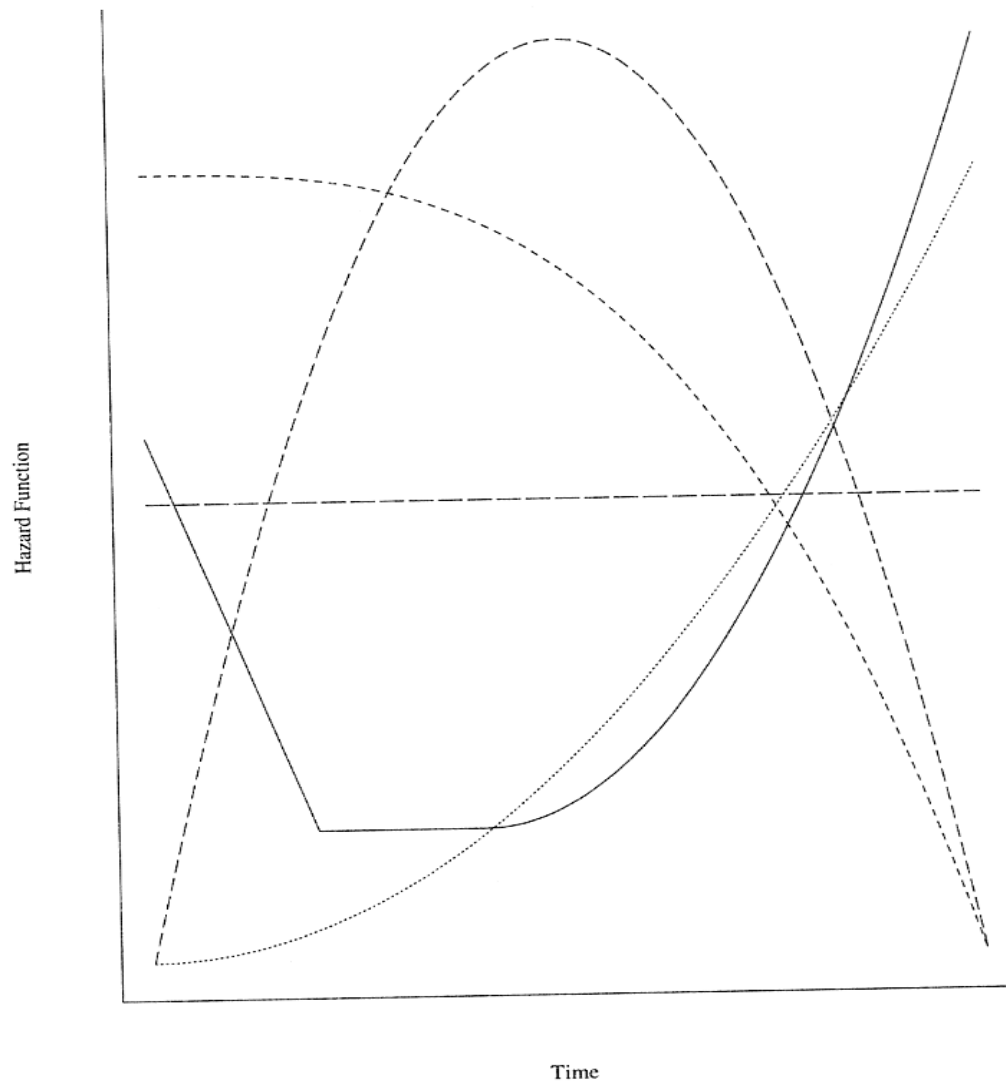


Figure 2.4 Shapes of hazard functions. Constant hazard (— — —); increasing hazard (· · · · ·); decreasing hazard (- - - - -); bathtub shaped (— — —); humpshaped (— — —).

Weibull-Verteilung

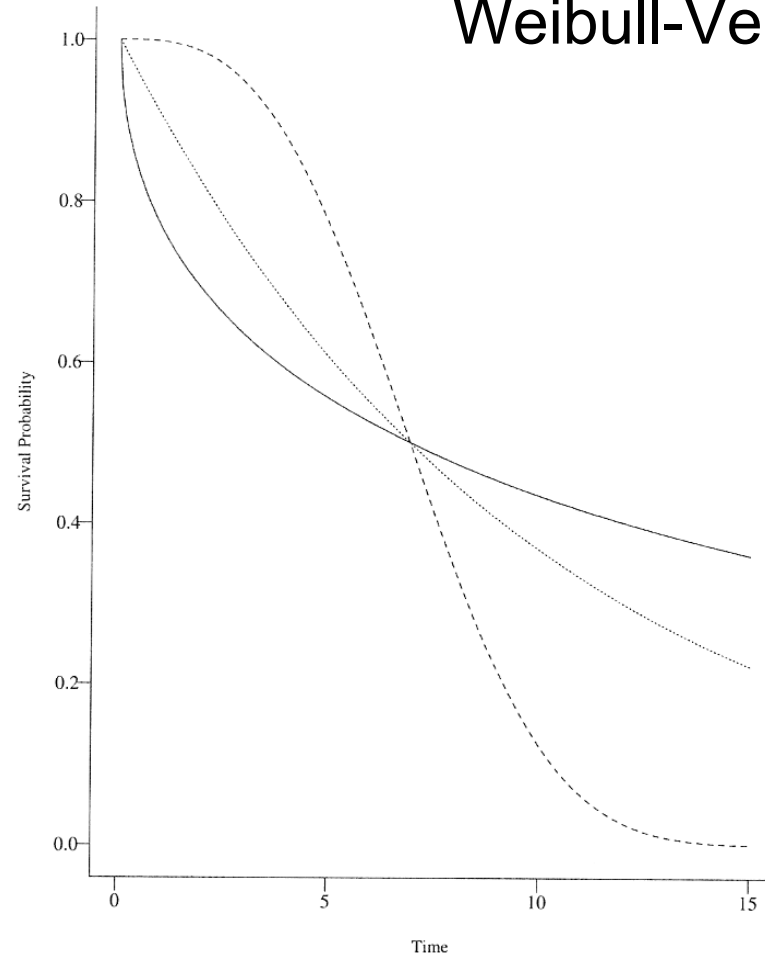


Figure 2.1 Weibull Survival functions for $\alpha = 0.5$, $\lambda = 0.26328$ (—); $\alpha = 1.0$, $\lambda = 0.1$ (·····); $\alpha = 3.0$, $\lambda = 0.00208$ (-----).

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

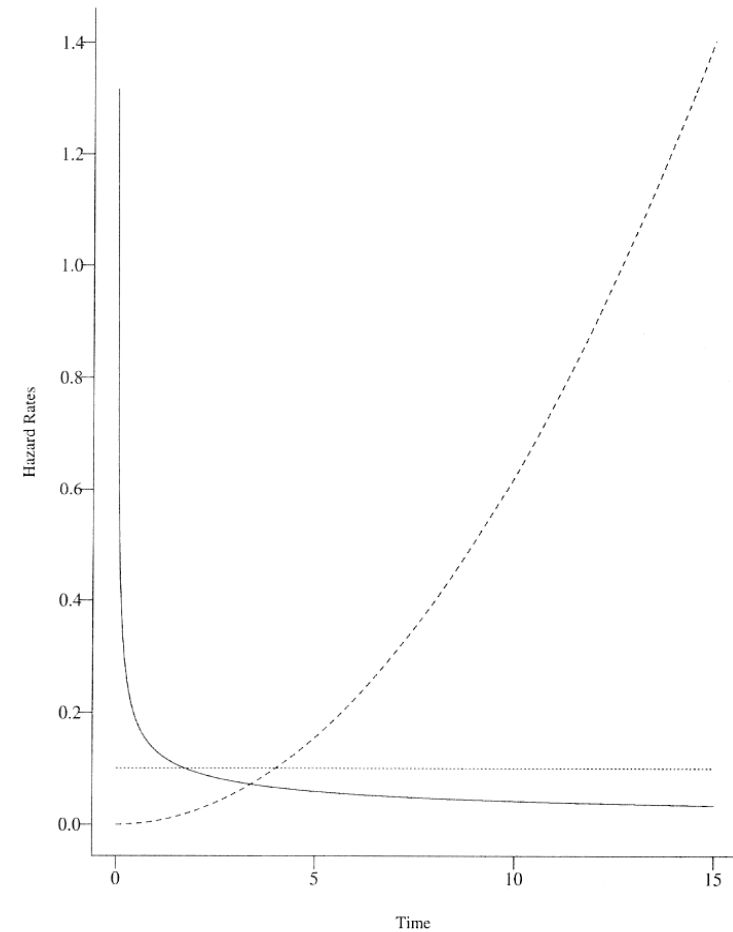


Figure 2.5 Weibull hazard functions for $\alpha = 0.5$, $\lambda = 0.26328$ (—); $\alpha = 1.0$, $\lambda = 0.1$ (-----); $\alpha = 3.0$, $\lambda = 0.00208$ (—··—).

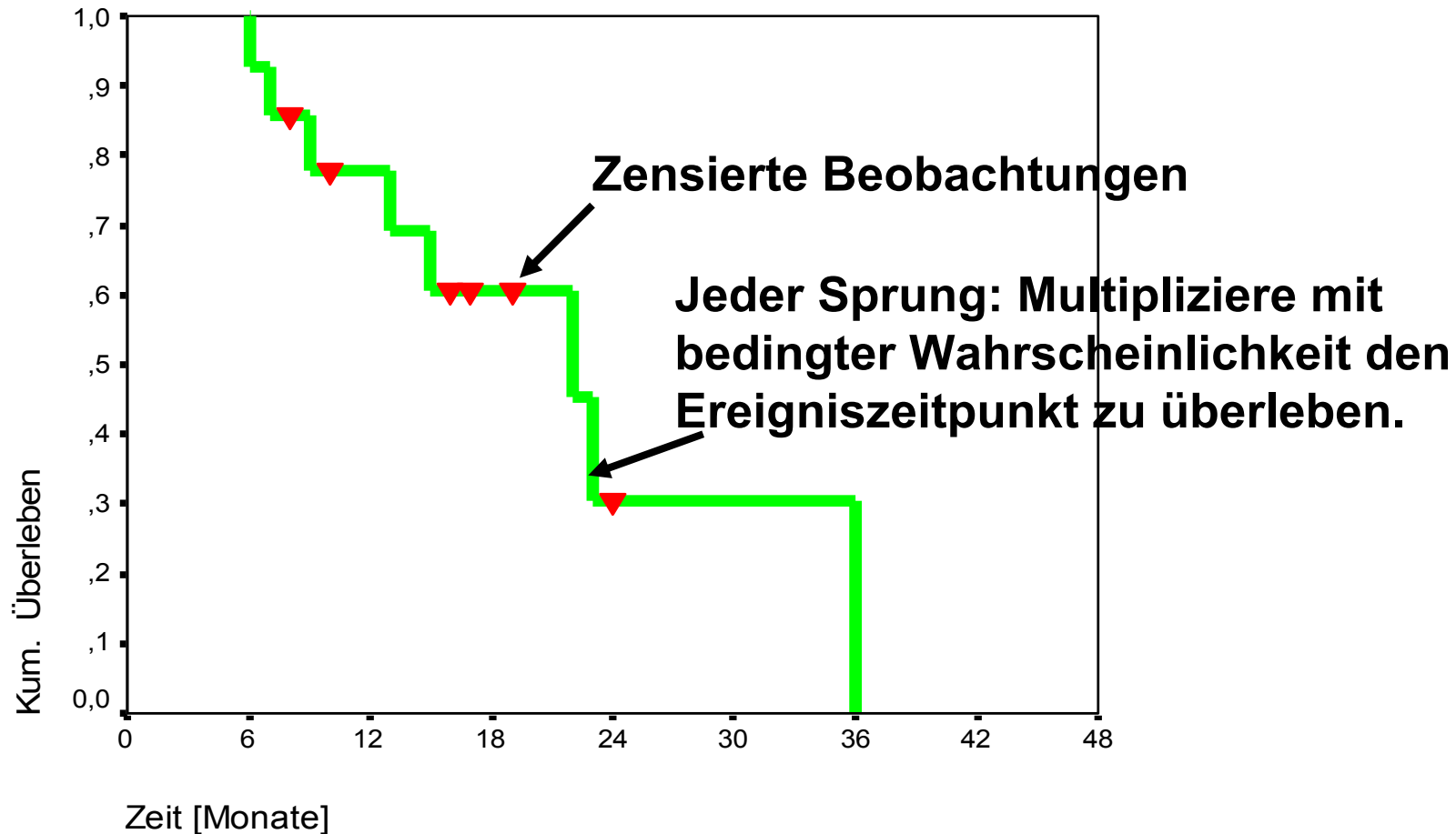
$$h(t) = \alpha \lambda t^{\alpha-1}$$

Grundidee: Kaplan Meier Schätzer I

- Bei Vorliegen von Zensur können Überlebenskurven nicht mehr durch einfache Häufigkeiten geschätzt werden.
- **Idee:** Um bis zum Zeitpunkt t zu überleben, muss ein Patient alle Zeitpunkte vorher, an denen jemand gestorben ist, überleben.
- Bedingte Wahrscheinlichkeit:

$$pr(T > t_j \mid T \geq t_j) \approx \frac{\# t_j - \text{überlebt}}{\# \text{bis} - t_j - \text{beobachtet}}$$

Grundidee: Kaplan Meier Schätzer II



„Um 100 zu werden, braucht man bloß das Jahr nach seinem 99. Geburtstag zu überleben.“

Grundidee: Kaplan Meier Schätzer III

Kaplan - Meier Schätzer $\equiv$ Product limit estimator

Aus den Daten bestimme die distinkten Ereigniszeitpunkte :

$$t_1 < t_2 < \dots < t_k$$

Sei $d_j = \# \text{Ereignisse zu } t_j$

$n_j = \# \text{Patienten, die zum Zeitpunkt } t_j \text{ unter Risiko stehen}$

$$\begin{aligned}\hat{S}(t) &= \prod_{t_j < t} \hat{\text{pr}}[T > t_j \mid T \geq t_j] = \prod_{t_j < t} \left(\frac{n_j - d_j}{n_j} \right) \quad (\text{Produkt über alle } j \text{ mit } t_j \leq t) \\ &= \prod_{t_j < t} \left(1 - \frac{d_j}{n_j} \right)\end{aligned}$$

Selbstkonsistenz des Kaplan Meier Schätzers

Sei $E(t) = \#$ Patienten, die mindestens bis Zeitpunkt t beobachtet wurden und bei denen bis zum Zeitpunkt t kein Ereignis beobachtet wurde.

Falls keine Zensur auftritt, so ist :

$$\hat{S}(t) = \frac{E(t)}{n}$$

der natürliche Schätzer für die invers kummulierte Verteilungsfunktion

Wenn ein Patient zum Zeitpunkt t_0 nicht - informativ (!) zensiert wurde, gilt für $t > t_0$:

$$pr(T > t \mid T > t_0) = \frac{S(t)}{S(t_0)}$$

Selbstkonsistenz des Kaplan Meier Schätzers II

Also sollte ein vernünftiger Schätzer folgende
(rekursive) Gleichung erfüllen :

$$\hat{S}(t) = \frac{E(t) + \sum_{\substack{\text{Patient} \\ \text{zensiert} \\ \text{in } t_j}} \frac{\hat{S}(t)}{\hat{S}(t_j)}}{n}$$

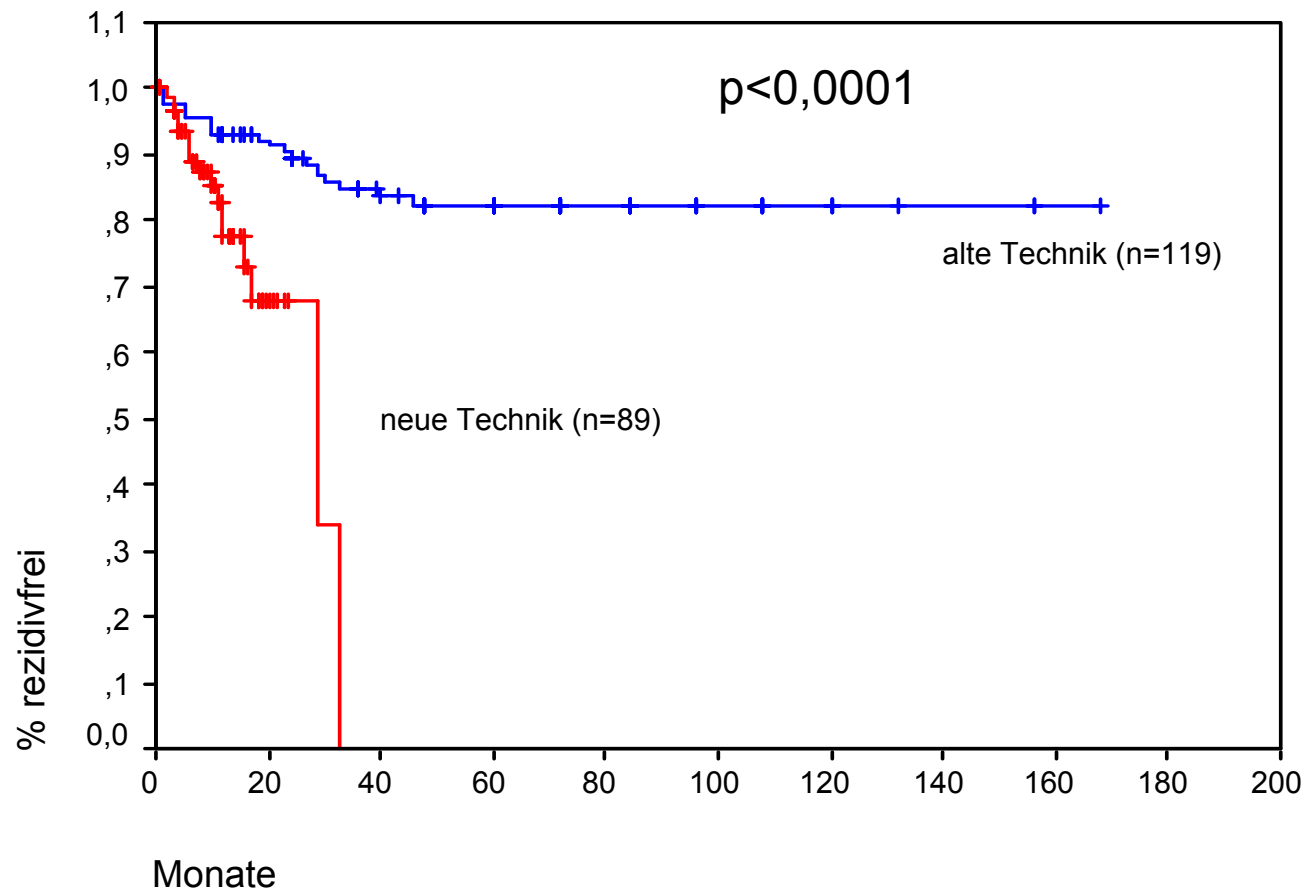
Man kann zeigen, dass das diese
Selbstkonsistenzbedingung
genau den product - limit - estimator charakterisiert.

„Zeit bis Ereignis“ – Daten IV

- **Vorsicht:** Gängige Auswertetechniken **nur** für **nicht-informative** Zensurmechanismen
- Grund für das Ende der Beobachtung muss **unabhängig vom Risiko für das Ereignis** sein (unter allen Patienten die mindestens bis zur gegebenen Zeit ohne Ereignis beobachtet wurden).
- Beispiel für informative Zensur: Ereignis=Tod. Patienten mit Rückfall der Erkrankung gehen zum Wunderheiler und entziehen sich systematisch der weiteren Beobachtung.

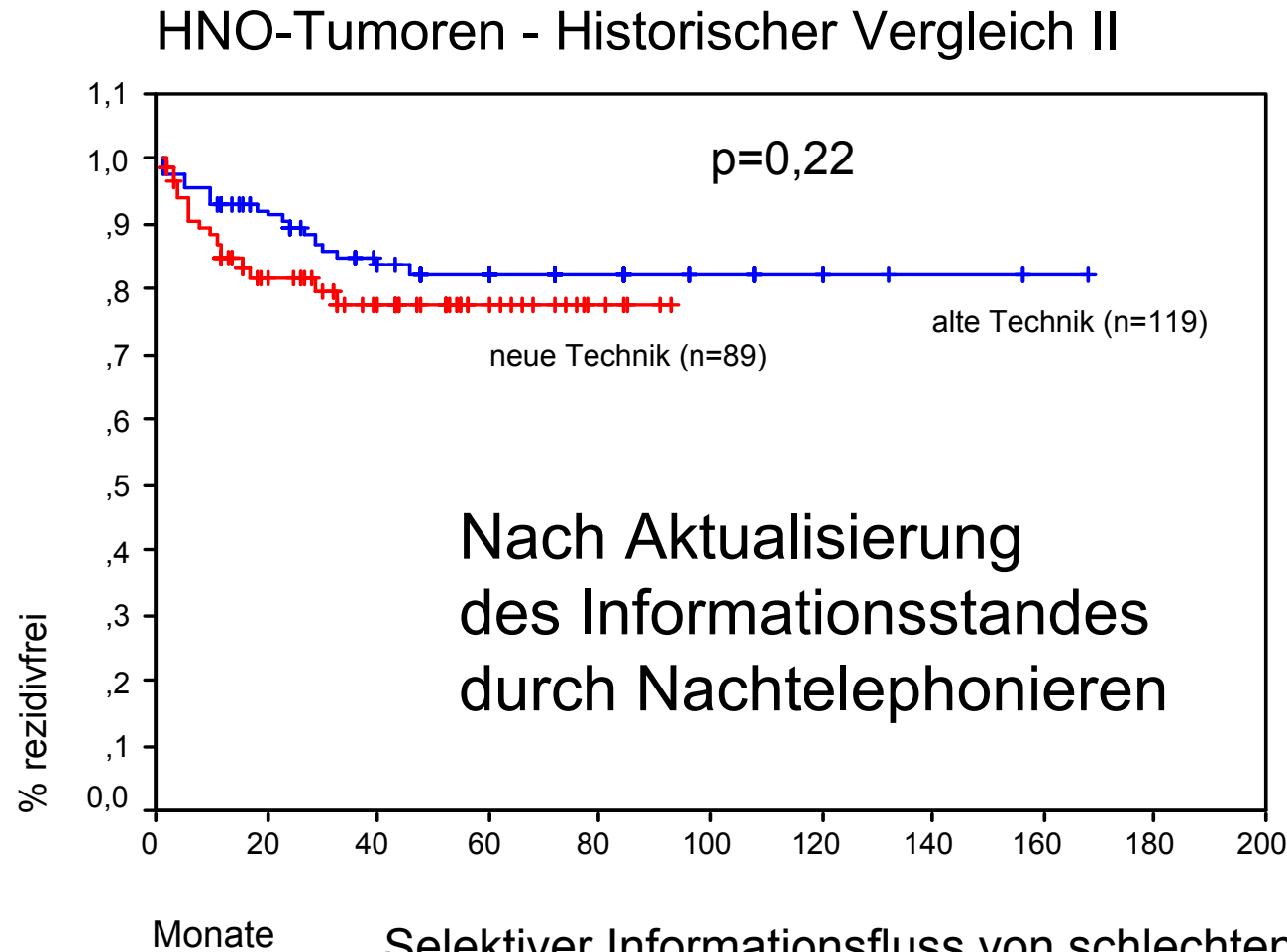
„Muß der Klinikscheff ins Gefängnis?“

HNO-Tumoren - Historischer Vergleich I



Vergleich der neuen
Bestrahlungs-Technik
bei HNO-Tumoren
mit der alten Technik
bzgl. rezidivfreien
Überlebens
nach Krankenaktenlage

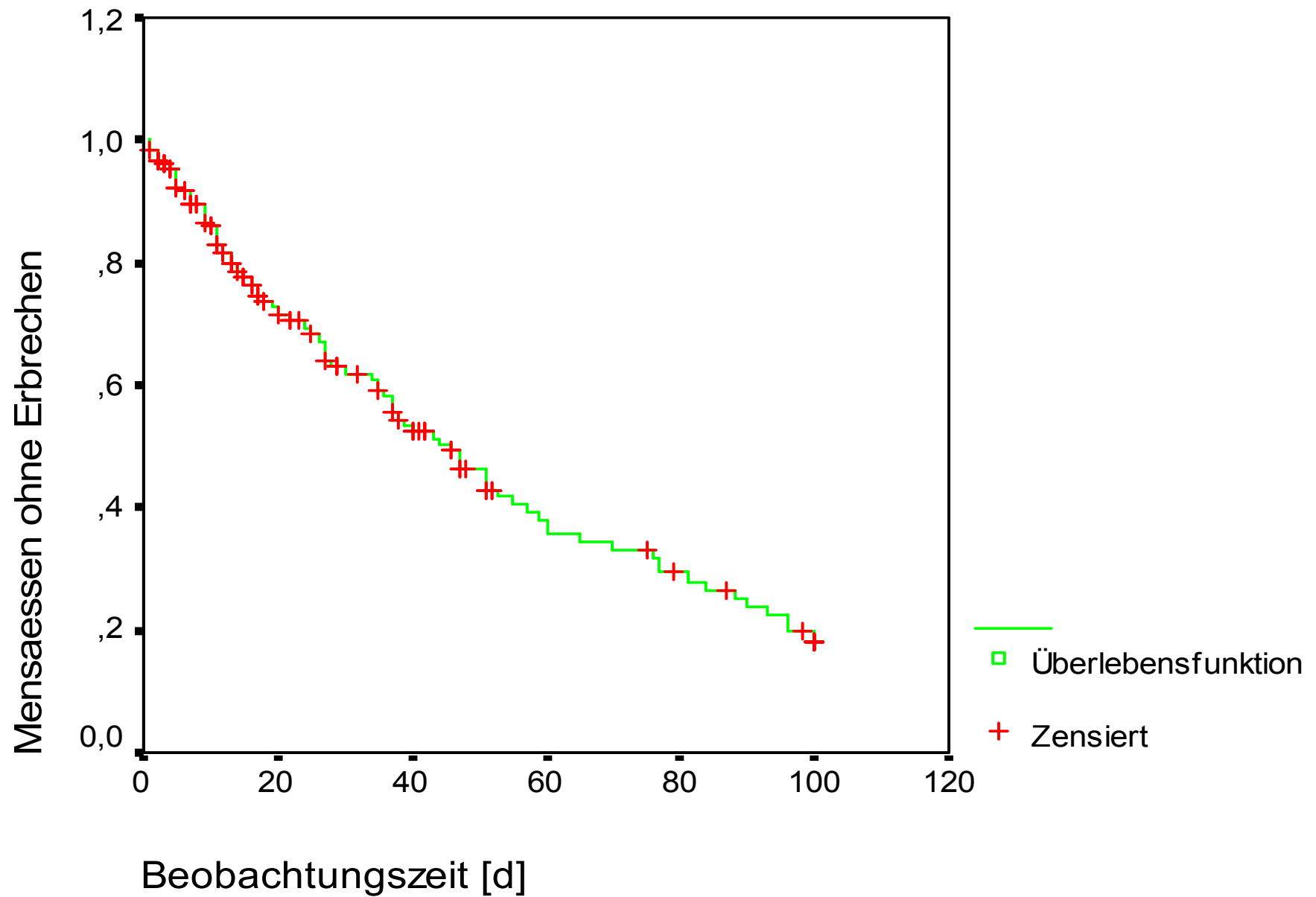
Informative Zensur – Zusammenhang zwischen Informationsfluss und Outcome



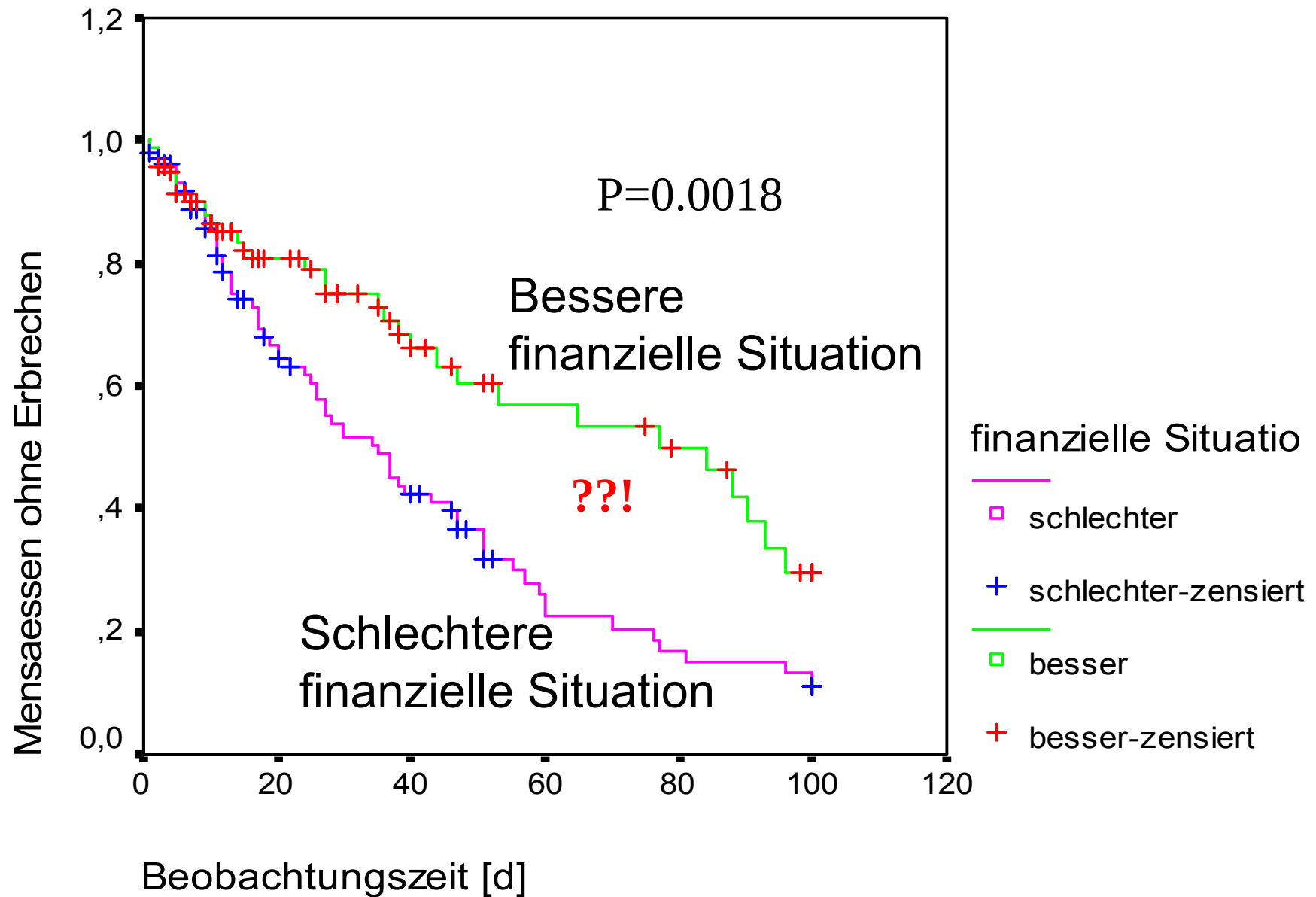
Studie zur Verträglichkeit des Mensaessens (Gelbrich)

- Randomisiert insgesamt N=200, motiviert durch Essenszuschuss
- Hauptendpunkt: Zeit bis Erbrechen
- Hypothese: “Reichere” Studenten sind verwöhnt und erbrechen daher schneller.
- Dokumentation der Abbruchgründe:
 - Studienende, Organisationsproblem, Ekel, Ereignis

Ereignis: Erbrechen



Ereignis = Erbrechen

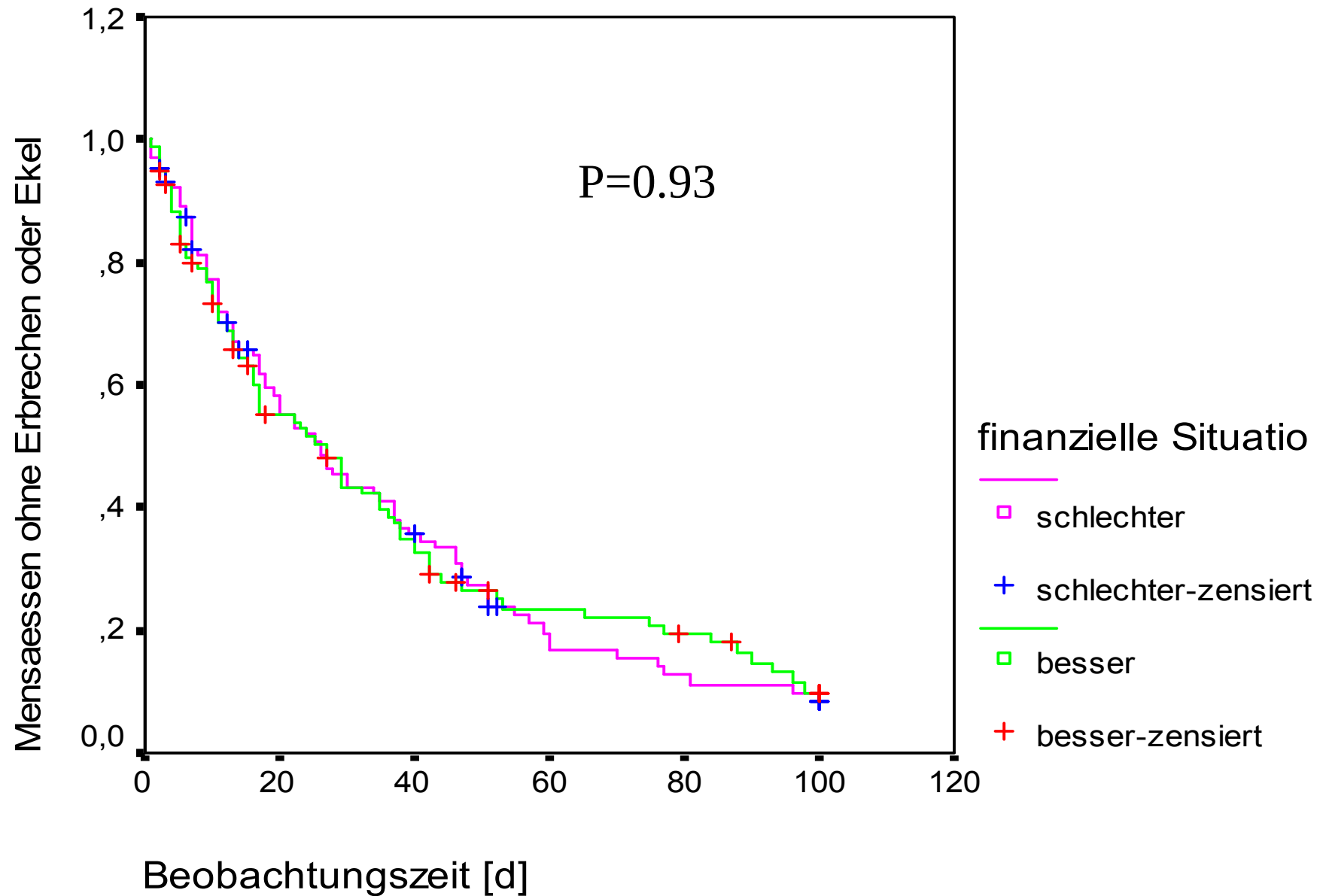


Studie zur Verträglichkeit des Mensaessens (Gelbrich)

Abbruchgrund * finanzielle Situation Kreuztabelle

			finanzielle Situation		Gesamt
			besser	schlechter	
Abbruchgrund	nicht zensiert	Anzahl	33	67	100
		% von finanzielle Situation	34,4%	64,4%	50,0%
	Ekel	Anzahl	43	16	59
		% von finanzielle Situation	44,8%	15,4%	29,5%
	organisatorisch	Anzahl	14	15	29
		% von finanzielle Situation	14,6%	14,4%	14,5%
	Studienende	Anzahl	6	6	12
		% von finanzielle Situation	6,3%	5,8%	6,0%
	Gesamt	Anzahl	96	104	200
		% von finanzielle Situation	100,0%	100,0%	100,0%

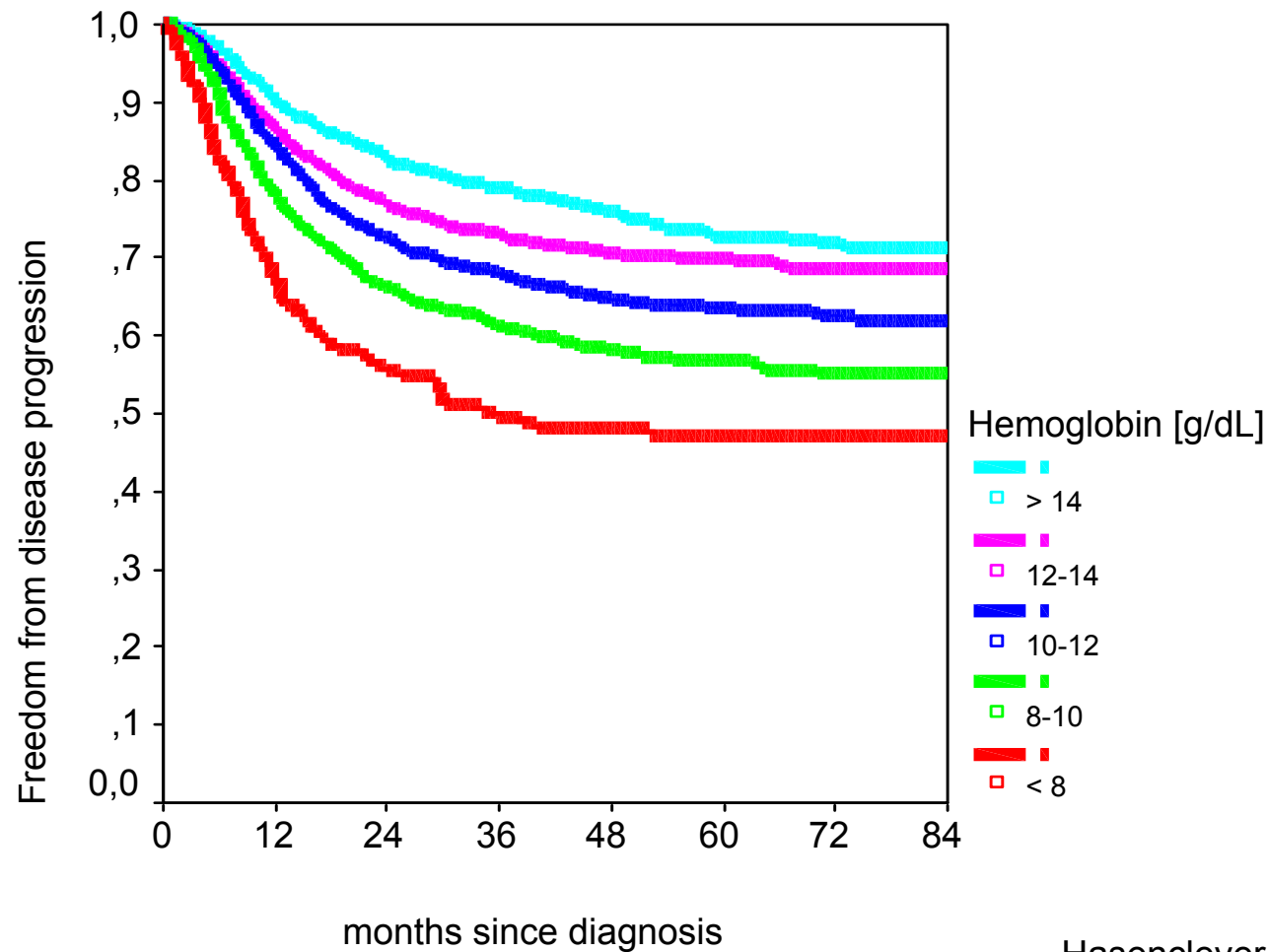
Ereignis = Erbrechen oder Ekel



Vorsicht: Informative Zensurmechanismen

- Methoden zum Schätzen von Überlebenskurven setzen in der Regel voraus, dass der Zensurmechanismus unabhängig von dem Ereignis-Hazard ist.
- D.h. Zensierte Patienten unterscheiden sich bzgl. des interessierenden Ereignisrisikos nicht, von weiterbeobachteten Patienten.
- Ist dies verletzt, können grobe Verzerrungen auftreten.

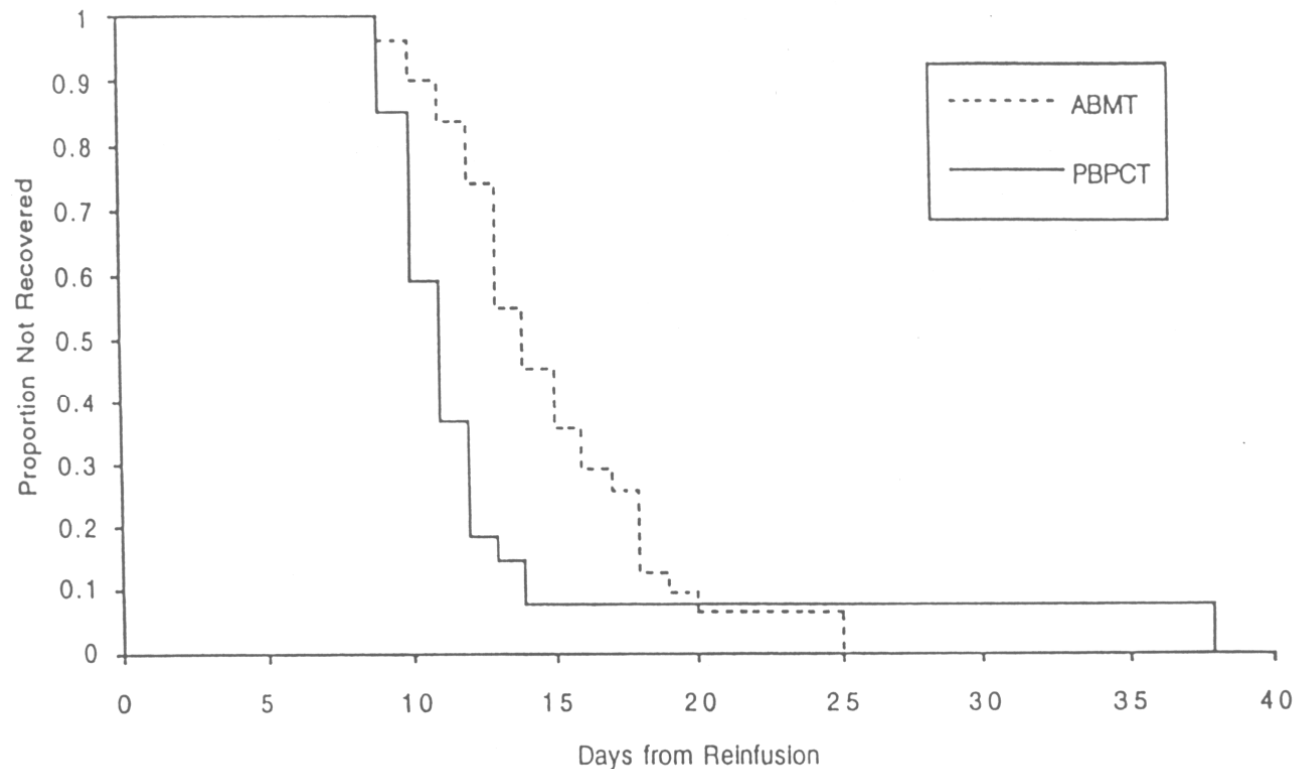
Kurventyp: Unterschiede im Plateau



Hasenclever et al. NEJM 1998

Kurventyp: Unterschied in medianer Zeit bis zum Ereignis

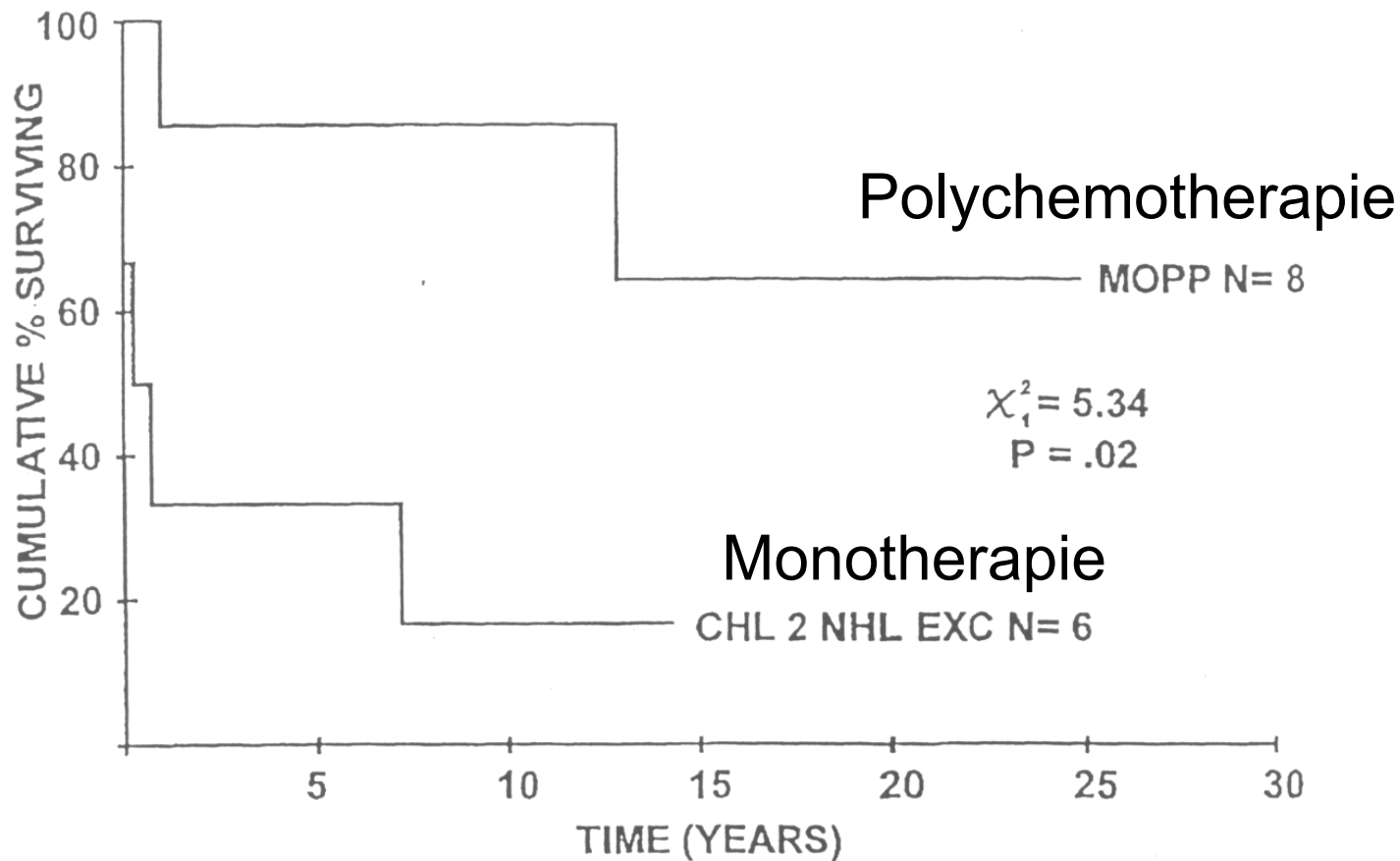
Zeit bis zur Erholung der Neutrophilen
nach autologer Stammzelltransplantation



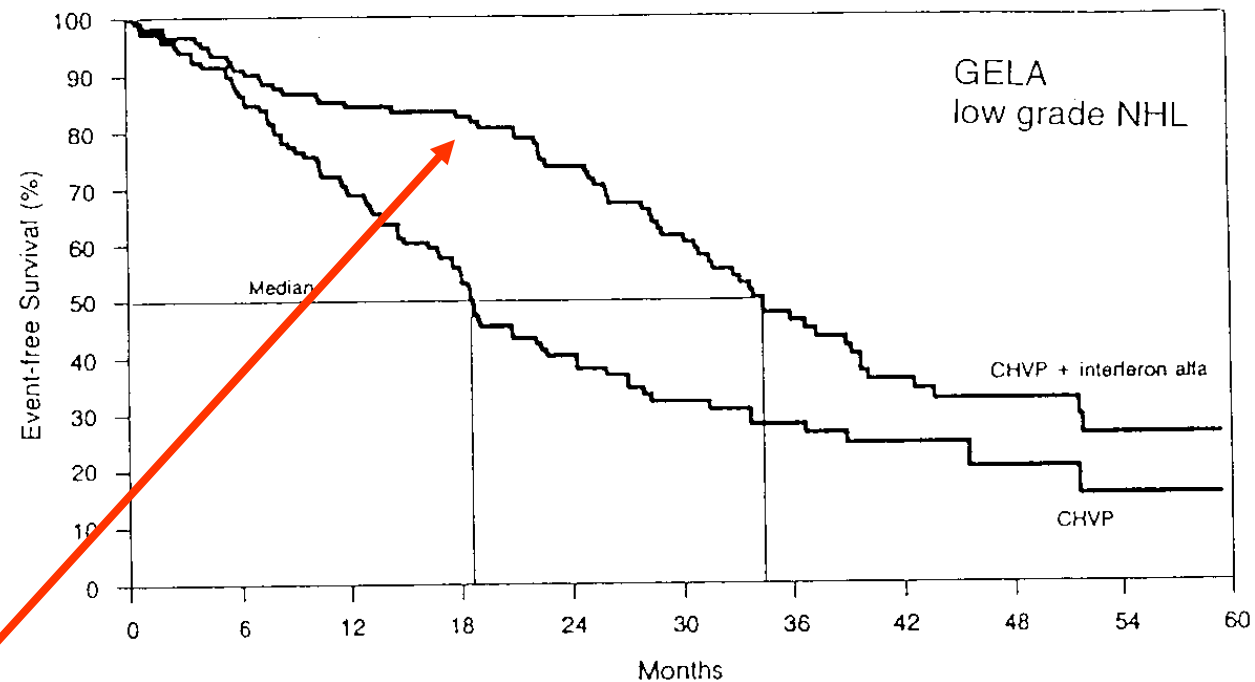
Schmitz et al. (Lancet 1996)

Kurventyp: Medizinischer Durchbruch (hier braucht man keine Statistik..)

Cause Specific Survival Stage 4 HD BNLI 1970



Kurventyp: Initialer Unterschied, der dann abklingt



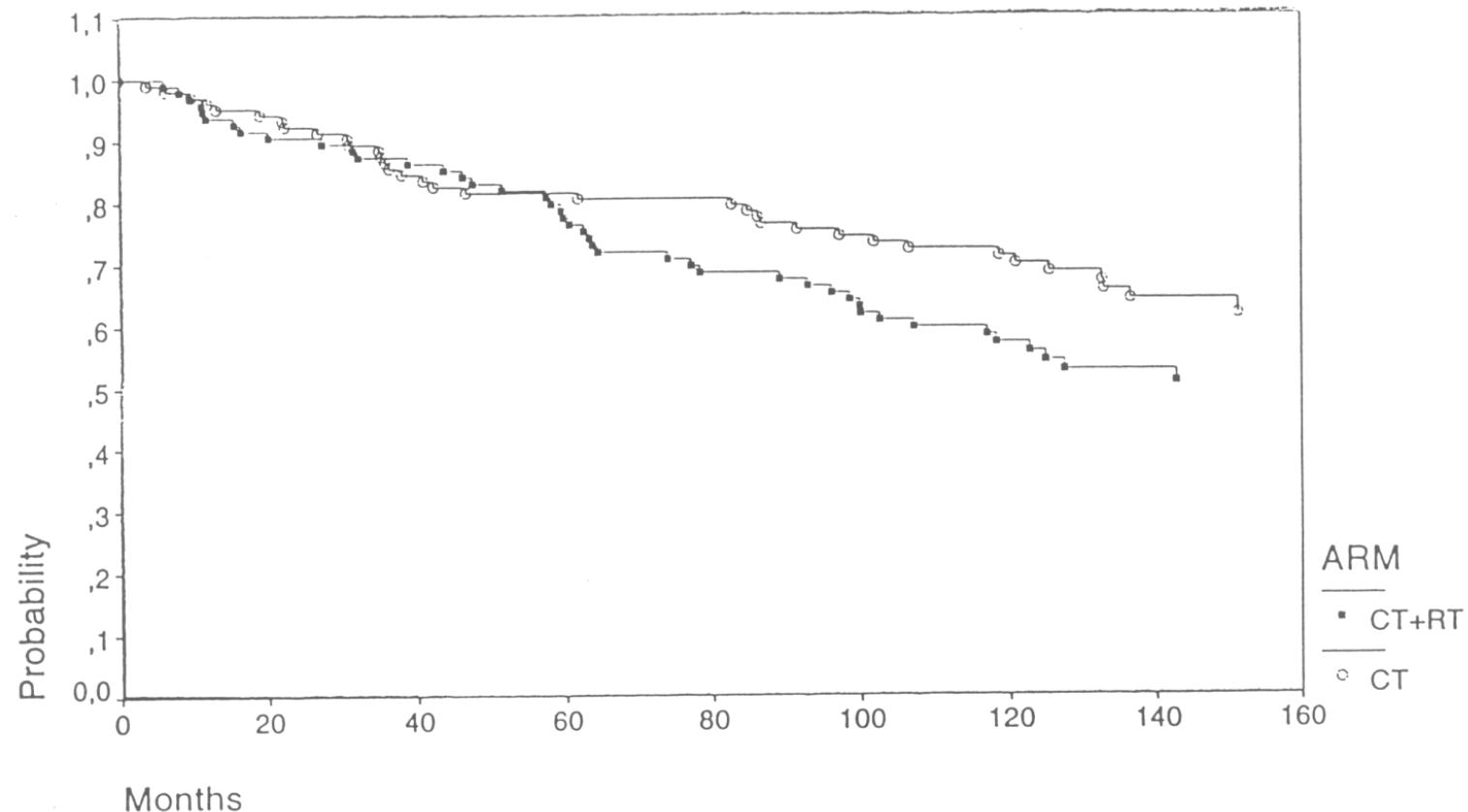
Ende der
Erhaltungstherapie
mit Interferon-alpha

Group	No. of Patients	No. with Events	Median Event-free Survival (95% CI)	P Value
CHVP	119	83	19 mo (17–23)	<0.001
CHVP + interferon alpha	123	65	34 mo (31–39)	

Figure 2. Estimated Event-free Survival According to Treatment Group.
CI denotes confidence interval.

Solal-Céligny et al. 1993, NEJM 329,22 1608-14

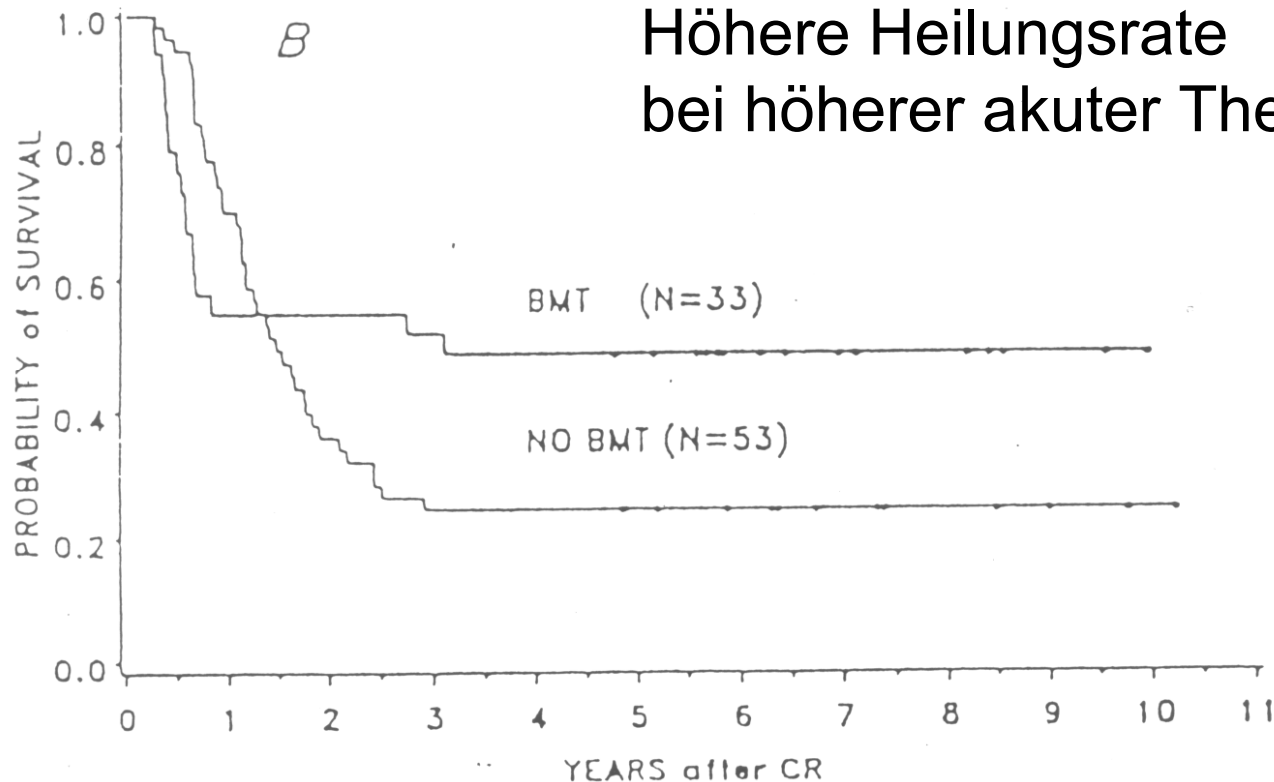
Identische Kurven, die nach einer Latenzzeit aufspreizen:



Unveröffentlichte Daten aus Meta-Analyse (Loeffler et al. (1998) JCO 16,3: 818-29)
Beim Hodgkin Lymphom: Chemotherapie allein gegen kombinierte Modalität

Sich kreuzende Vergleichskurven

Allogene Transplantation bei AML
Höhere Heilungsrate
bei höherer akuter Therapiemortalität.



Appelbaum et al. 1988, Blood 72,1:179ff