Inhaltsverzeichnis

Statistikfortbildung

Auswertung von Überlebensdaten in klinischen Studien der pädiatrischen Onkologie

Oktober 2001

Kompetenznetz Päd. Onkologie Institut für Biometrie, MHH

H.Hecker

1. Besonderheiten bei der Auswertung von Überlebensdaten

1.1. Was sind Überlebensdaten?

Bei der Auswertung von Überlebensdaten geht es allgemein um die Fragestellung:

> Wie lange dauert es, bis ein definiertes Ereignis (Progression, Relapse, Tod) eingetritt?

Bei jedem Patienten werden die Zeitpunkte

- Beginn der Beobachtung
- Ende der Beobachtung

und der Status bei Ende der Beobachtung:

- 0 = Ereignis ist nicht eingetreten (unvollständige Beobachtung: "zensiert")
- 1 = Ereignis ist zu diesem Zeitpunkt eingetreten

notiert.
Graphische Darstellung:

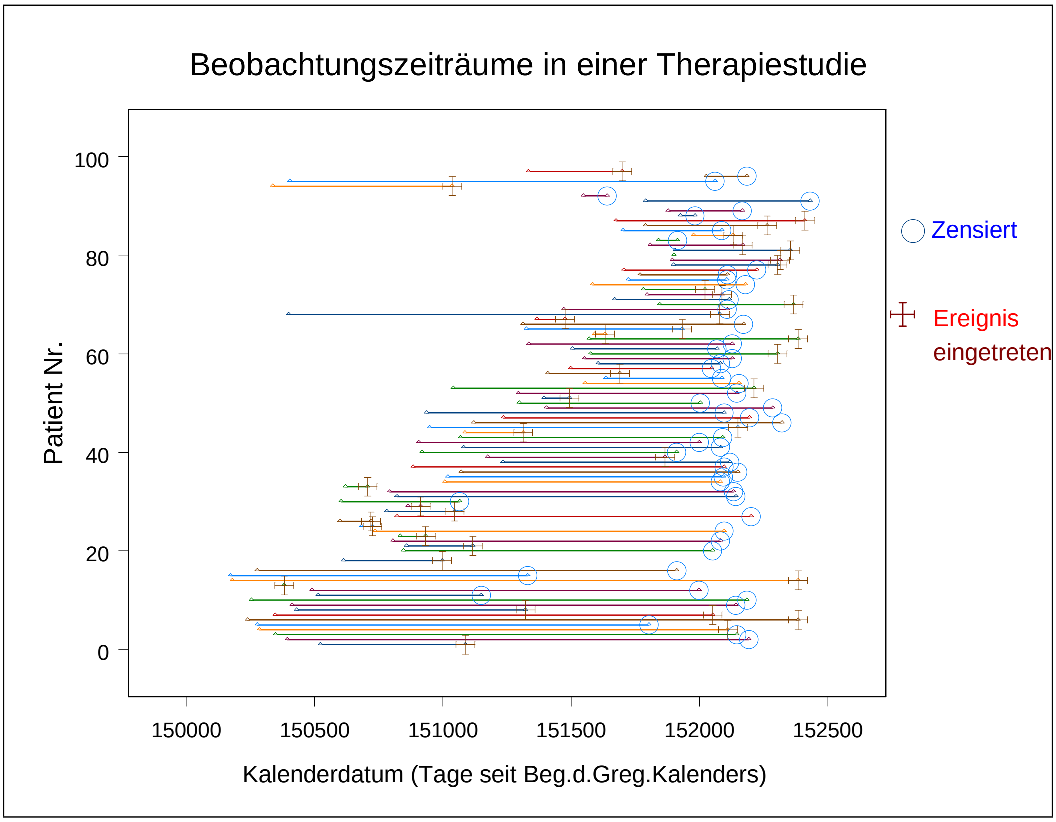

Abbildung 1

Da man sich für die *Dauer* der Beobachtung interessiert (unabhängig davon, wann sie begonnen hat), wird der Beginn für jeden Patienten auf Null gesetzt:

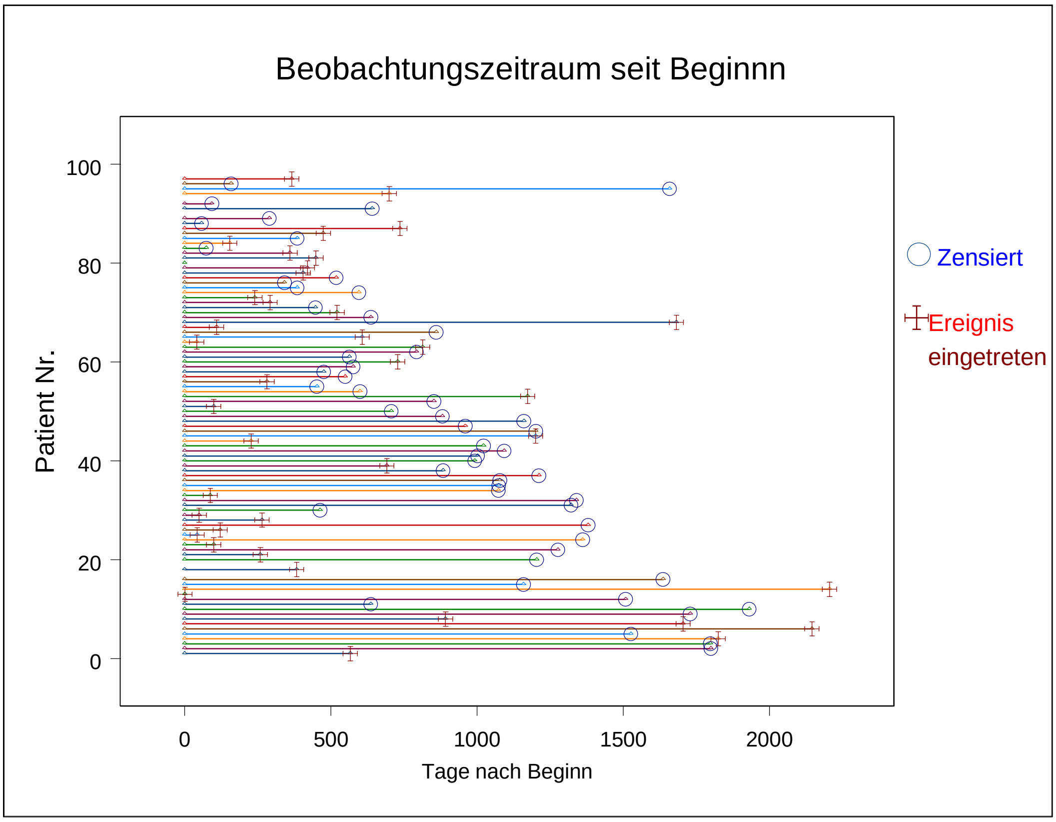

Abbildung 2

1.2. Wie werden die Daten gespeichert?

Man bildet eine *Rechteckdatei*, in der jede *Zeile* die Daten eines *Patienten* enthält und die einzelnen *Spalten* für die verschiedenen *Variablen* reserviert sind:

	patient	beginn	ende	dauer	statusr
1	1	150522	151088	566	1
2	2	150393	152192	1799	0
3	3	150348	152146	1798	0
4	4	150286	152110	1824	1
5	5	150277	151803	1526	0
6	6	150239	152384	2145	1
7	7	150347	152051	1704	1
8	8	150430	151322	892	1
9	9	150413	152142	1729	0
10	10	150254	152185	1931	0
11	11	150514	151150	636	0
12	12	150490	151998	1508	0
13	13	150381	150382	1	1
14	14	150179	152384	2205	1
15	15	150172	151331	1159	0

Abbildung 3

Welches sind nun die Überlebenszeiten der einzelnen Patienten in diesem Beispiel?
Patient 1: 566 Tage
Patient 2: Weiß man nicht genau, jedenfalls länger als 1799 Tage!

Will man also noch eine Spalte "Überlebenszeit" einführen, so muss diese unvollständig sein:

	patient	beginn	ende	dauer	statusr	survival
2	2	150393	152192	1799	0	? (> 1799)
3	3	150348	152146	1798	0	? (> 1799
4	4	150286	152110	1824	1	1824
5	5	150277	151803	1526	0	? (> 1526)
6	6	150239	152384	2145	1	2145
7	7	150347	152051	1704	1	1705
8	8	150430	151322	892	1	892
9	9	150413	152142	1729	0	? (> 1729)
10	10	150254	152185	1931	0	? (> 1931)
11	11	150514	151150	636	0	? (>636)
12	12	150490	151998	1508	0	? (> 1508)
13	13	150381	150382	1	1	1
14	14	150179	152384	2205	1	2205
15	15	150172	151331	1159	0	? (> 1159)

Abbildung 4

(Hinweis: Mit der Spalte "survival" kann man in der Auswertung nichts anfangen. Sie ist hier nur zur Demonstration angelegt!)

Wie groß ist nun z.B. die *mittlere Überlebenszeit*?
Kann man nicht berechnen, da die Daten unvollständig sind!

Wie wäre es,

- den Mittelwert über alle *Beobachtungszeiten* (Spalte 4: "dauer") zu berechnen
  - der wäre zu klein! (Begründung?)
- die Patienten mit unvollständigen Daten wegzulassen?
  - der MIttelwert könnte zu groß oder zu klein sein, je nachdem ob die zensierten Patienten im Schnitt länger oder kürzer leben.

Was sind überhaupt die Gründe für "Zensierung"?

- "Lost to follow up", "withdrawn from follow up": keine weitere Information erhältlich, weil andere Ereignisse dazwischengekommen sind.
- das Ereignis ist bis zum Ende der Beobachtungsphase (bis zum Ende der Studie) nicht eingetreten. Hiervon sind primär betroffen:
  - Patienten mit spätem Eintritt in die Studie;
  - Patienten mit besonders langem Überleben.

Die letzte genannte Ursache würde eher zu einer Unterschätzung der mittleren Überlebenszeit führen, wenn man zensierte Fälle wegließe.

Anderer Versuch zur statistischen Aufarbeitung:

Wie groß ist die Überlebenswahrscheinlichkeit nach 5 Jahren?

Man notiert dazu bei jedem Patienten, ob er 5 Jahre (=1826 Tage) überlebt hat. Ergebnis in der folgenden Abbildung:

muster0.sav - SPSS Data Editor

	patient	beginn	ende	dauer	statusr	survival	s_5_jahr
1	1	150522	151088	566	1	566	2
2	2	150393	152192	1799	0	? (> 1799)	3
3	3	150348	152146	1798	0	? (> 1798)	3
4	4	150286	152110	1824	1	1824	2
5	5	150277	151803	1526	0	? (> 1526)	3
6	6	150239	152384	2145	1	2145	1
7	7	150347	152051	1704	1	1705	2
8	8	150430	151322	892	1	892	2
9	9	150413	152142	1729	0	? (> 1729)	3
10	10	150254	152185	1931	0	? (> 1931)	1
11	11	150514	151150	636	0	? (>636)	3
12	12	150490	151998	1508	0	? (> 1508)	3
13	13	150381	150382	1	1	1	2
14	14	150179	152384	2205	1	2205	1
15	15	150172	151331	1159	0	? (> 1159)	3

Abbildung 5

Darin ist die letzte Spalte ("s-5_jahr": Survival nach 5 Jahren) wie folgt kodiert:

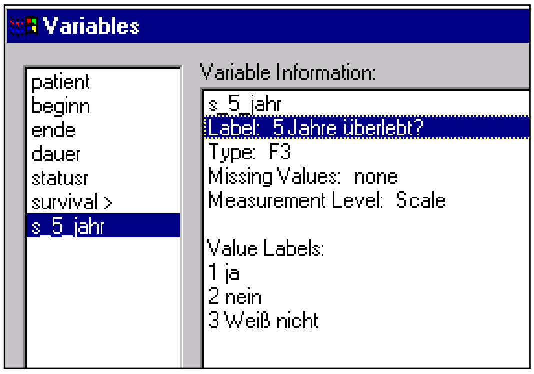

Abbildung 6

Man sieht:

Zensierte Fälle *können* immer noch ein Problem darstellen, müssen dies aber nicht mehr notwendigerweise. Beispielsweise ist die Beobachtung von Patient Nr. 10 "zensiert"

nach 1931 Tagen, aber zur Frage nach der 5-Jahres-Überlebensrate ist die Information bei ihm vollständig: Er hat die ersten 5 Jahre (=1826 Tage) nach Beginn überlebt.
Noch günstiger wird die Situation, wenn man z.B. nach der 1-Jahres-Überlebenswahrscheinlichkeit fragt. Aus Abbildung 2 sieht man, dass dann nur noch *sehr wenige* Patienten *unvollständige Information* zu dieser Frage liefern, nämlich alle "Früh-Zensierten" (Zensierung im ersten Jahr).
Dennoch bleibt das Problem der unvollständigen Information gegeben. Soll man die zensierten Fälle aus der Ratenberechnung weglassen? Das Ergebnis wäre unsicher, da man nicht weiß, wie groß die Überlebensrate dieser "Früh-Zensierten" im Verleich zu den anderen Patienten wäre.

2. Berechnung von Überlebenswahrscheinlichkeiten

2.1. Lifetable Methode

Man muss nach dem letzten Ansatz nur noch eine kleine Überlegung anstellen, um zum ersten Schritt der "Life Table Methode" zu kommen:

1. Bilde das erste Zeitintervall (z.B. Beginn bis 1 Jahr nach Beginn).
2. Zähle die in diesem Intervall "zensierte" Fälle.
3. Gehe davon aus, dass diese Patienten im Schnitt das halbe Jahr überlebt haben.
4. Bewerte *zwei Patienten mit jeweils einem halben* überlebten Intervall wie *einen* Patienten mit einem *ganzen* überlebten Intervall.
5. Dann ist schon alles klar!

Nämlich:

$$\boxed{\text{Sterberate im 1. Interv.} = \frac{\text{Anzahl Verstorbener im 1. Intervall}}{\text{Anz. Pat. zu Beginn} - 1/2 \times \text{Anz. Zens im 1.Interv.}}}$$

$$\boxed{\text{Überlebensrate im 1. Intervall} = 1 - \text{Sterberate im 1. Intervall}}$$

Solche Sterbe- und Überlebensraten kann man auch für andere Zeitintervalle bilden. Dann fragt man z.B. nach der Überlebensrate für das 2. Jahr von allen Patienten, *die zu Beginn des ersten Jahres noch in der Beobachtung sind* usw. Das Berechnungsschema dafür kann man jetzt schon aufschreiben:

Spalte- Nr. Inter- vall Nr.	S1 Lebend zu Beginn	S2 Zensiert im Intervall	S3 Anzahl unter Risiko $(S1 - 1/2 \times S2)$	S4 Verstorben im Intervall	S5 Sterbe- rate $\frac{S4}{S3}$	S6 Überl.Rate im Intervall $1 - S5$

Aufgabe:
Ergänze die erste Zeile dieser Tabelle nach den Daten der Abbildung 6 (1.Intervall = Beginn bis 5 Jahre).

Die größte Unsicherheit bei dieser Berechnung liegt in der Behandlung der Daten, die im Intervall zensiert wurden:

1. ist die Annahme von durchschnittlich einem halbem Intervall überlebter Zeit je Patient unsicher
2. könnte das Sterberisiko innerhalb des Intervalls stärker variieren, so dass das "Aneinandersetzen" von beobachteten Überlebenszeiten nicht gerechtfertigt erscheint.

Um diese Unsicherheit aufzufangen, bildet man möglichst kleine Intervalle. Um trotzdem den maximalen Beobachtungszeitraum zu erfassen, muss man die Information aus den einzelnen Intervallen **zusammensetzen** Das ist der eigentliche Trick der " Lifetable Methode" und geht nach folgender einfacher Überlegung:

Frage:
Wie groß ist die Überlebenswahrscheinlichkeit nach dem 2., 3., 4., … Intervall?

Hypothetisches Beispiel:

- Das 1. Intervall überleben 70 %
- Das 2. Intervall überleben 80 % *der Patienten*, die das 1. Intervall überlebt haben
- Das 3. Intervall überleben 90 % *der Patienten*, die das 2. Intervall überlebt haben

usw.
Dann ist die Anzahl der Überlebenden von zunächst 100 Patienten

- nach dem 1. Intervall: $\boxed{100 \times 0.7 = 70}$
- nach dem 2. Intervall: $\boxed{70 \times 0.8 = 56}$
- nach dem 3. Intervall: $\boxed{56 \times 0.9 = 50.4}$

usw.
Nun können wir die Tabelle zur Berechnung der Überlebensraten *insgesamt* (nicht nur für die einzelnen Intervalle) für die ganze Beobachtungszeit ergänzen:

	$S1$	$S2$	$S3$	$S4$	$S5$	$S6$	$S7$
Intervall Nr.	Lebend zu Beginn	Zensiert im Intervall	Anzahl unter Risiko $(S1 - 1/2 \times S2)$	Verst. im Interv.	Sterberate $\frac{S4}{S3}$	Überl.Rate im Intervall $1 - S5$	Gesamt-Überlebens-Rate $[S7(i-1)] \times S6$
1							
2							
3							

Darin bedeutet $[S7(i-1)]$ die Überlebensrate aus dem jeweils *vorangehenden Intervall* (Nr. $(i-1)$). Sie beginnt mit $1 = 100$ %.

Beispiel 1

Aufgabe:

Ergänze die folgende Lifetable (Daten aus: "M. Bland: An introduction to medical statistics"):

Inter-vall Nr.	$S1$ Lebend zu Beginn	$S2$ Zensiert im Intervall	$S3$ Anzahl unter Risiko $(S1 - 1/2 \times S2)$	$S4$ Verst. im Interv.	$S5$ Sterbe-rate $\frac{S4}{S3}$	$S6$ Überl.Rate im Intervall $1 - S5$	$S7$ Gesamt-Überlebens-Rate $[S7(i-1)] \times S6$
1	20	2		1			
2	17	2		0			
3	15	0		1			
4	14	0		0			
5	14	1		0			
6	13	1		0			
7	12	1		2			
8	9	0		1			
9	8	1		0			
10	7	0		2			
11	5	2		0			
12	3	0		1			
13	2	0		0			
14	2	0		0			
15	2	0		1			
16	1	0		0			
17	1	0		0			
18	1	1		0			

Beispiel 2

Berechnung der Lifetable zu den Daten der Abbildung 2 mit Hilfe von SPSS

This subfile contains: 92 observations

Life Table
Survival Variable DAUER Dauer der Beobachtung

Intrvl Start Time	Number Entrng this Intrvl	Number Wdrawn During Intrvl	Number Exposd to Risk	Number of Termnl Events	Propn Termi-nating	Propn Sur-viving	Cumul Propn Surv at End
.0	92.0	6.0	89.0	17.0	.1910	.8090	.8090
365.3	69.0	16.0	61.0	12.0	.1967	.8033	.6498
730.5	41.0	13.0	34.5	3.0	.0870	.9130	.5933
1095.8	25.0	10.0	20.0	2.0	.1000	.9000	.5340
1461.0	13.0	7.0	9.5	3.0	.3158	.6842	.3654
1826.3	3.0	1.0	2.5	1.0	.4000	.6000	.2192
2191.5	1.0	.0	1.0	1.0	1.0000	.0000	.0000

The median survival time for these data is 1534.6

Graphische Abbildung dazu (ohne weitere "Aufbereitung"):

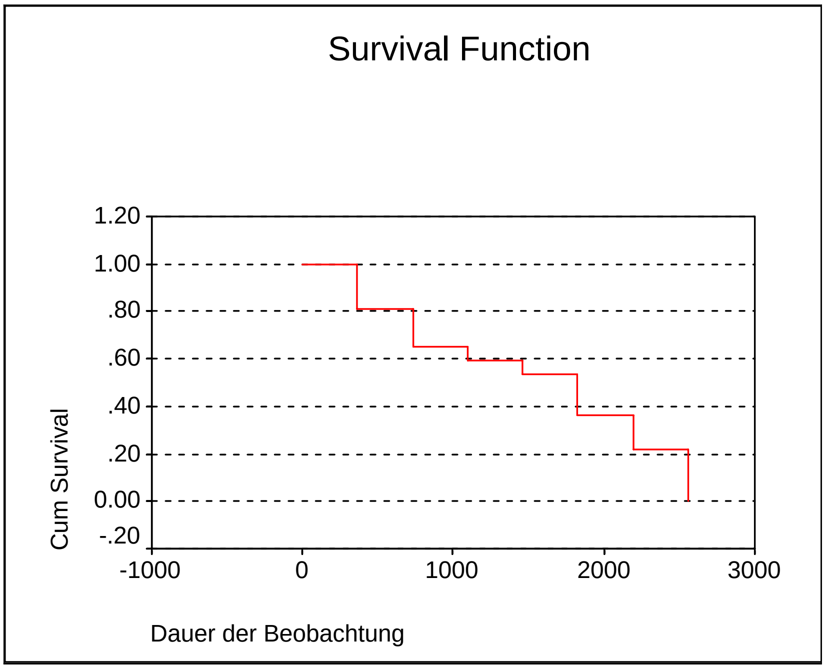

Abbildung 7

Beispiel 3.

Berechnung der Lifetable zu den Daten der Abbildung 2 mit Hilfe von *ganz kleinen Intervallen*

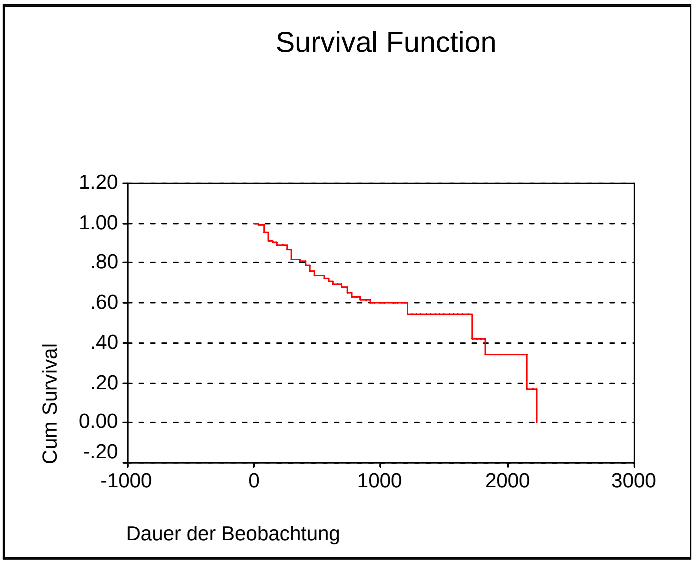

Abbildung 8

2.2. Methode nach Kaplan-Meier

1. Voraussetzung: Die Beobachtungszeiten sind alle exakt bekannt (man kennt also nicht nur den Zustand jeweils zum Ende einese Intervalls).

2. Wähle in der Lifetable Methode die Intervalle wirklich **sehr klein!**

3. Dann gibt es viele Intervalle, in denen gar ichts passiert; die kann man in der Berechnung gleich weglassen.

4. Man überlege: Wie ändert sich die Lifetable in den Intervallen, in denen nur zensiert wird und keine "Fälle" eintreten? (Antwort: Die Überlebenswahrscheinlichkeit ändert sich nicht, nur die Anzahl Patienten "at risk" wird für das nächste Intervall kleiner.)

5. Man überlege: Wie ändert sich die Lifetable in den Intervallen, in denen nur 1 Sterbefall eintritt? (Antwort: Die Überlebenswahrscheinlichkeit wird um den Faktor $1/n_i$ kleiner, wenn n_i die Anzahl Patienten zu Beginn des Intervalls bezeichnet.)

Bei der Methode nach Kaplan-Meier werden diese Überlegungen systematisch durchgeführt und die entsprechenden Formeln daraus abgeleitet. Sie resultiert als Grenzfall "beliebig kleiner Intervalle" und wird deshalb auch "product limit" Methode genannt.

Beispiel
Es werden wieder die Daten aus Abbildung 2 benutzt.
Ergebnis:

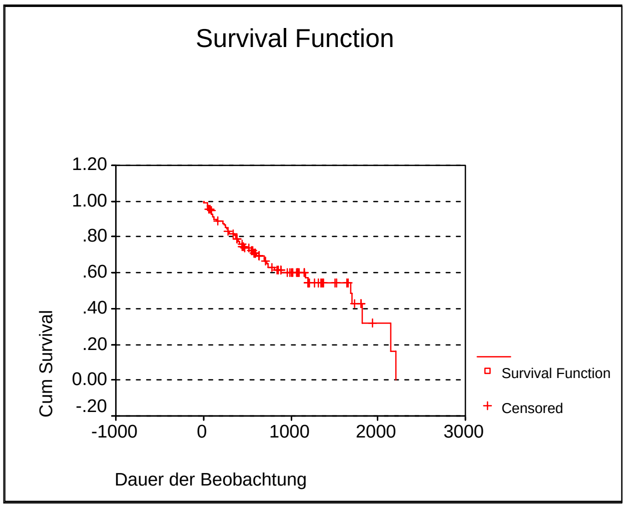

Abbildung 9

und der erste Teil der Tabelle:

Survival Analysis for DAUER   Dauer der Beobachtung

Time	Status	Cumulative Survival	Standard Error	Cumulative Events	Number Remaining
1	Rezidiv - Tod	.9891	.0108	1	91
41	Rezidiv - Tod	.9783	.0152	2	90
42	Rezidiv - Tod	.9674	.0185	3	89
49	Rezidiv - Tod	.9565	.0213	4	88
58	Censored - rezidivfrei			4	87
74	Censored - rezidivfrei			4	86
87	Rezidiv - Tod	.9454	.0237	5	85
93	Censored - rezidivfrei			5	84
99	Rezidiv - Tod			6	83
99	Rezidiv - Tod	.9229	.0280	7	82
109	Rezidiv - Tod	.9116	.0298	8	81
121	Rezidiv - Tod	.9004	.0315	9	80
154	Rezidiv - Tod	.8891	.0331	10	79
159	Censored - rezidivfrei			10	78
227	Rezidiv - Tod	.8777	.0346	11	77
240	Rezidiv - Tod	.8663	.0359	12	76
258	Rezidiv - Tod	.8549	.0372	13	75
264	Rezidiv - Tod	.8435	.0384	14	74
281	Rezidiv - Tod	.8321	.0396	15	73
290	Censored - rezidivfrei			15	72
292	Rezidiv - Tod	.8206	.0407	16	71
341	Censored - rezidivfrei			16	70
360	Rezidiv - Tod	.8089	.0418	17	69
366	Rezidiv - Tod	.7971	.0428	18	68
383	Rezidiv - Tod	.7854	.0437	19	67
385	Censored - rezidivfrei			19	66
385	Censored - rezidivfrei			19	65

Interpretation:

1. Die Überlebenskurve ist ähnlich derjenigen in Abbildung 9, die als Lifetable mit kleinenr Intervallen erzeugt war.

2. Man kann für jeden Zeitpunkt aus der Graphik bzw. aus der Tabelle die Überlebenswahrscheinlichkeit ablesen.

3. Man kann ablesen, zu welchem Zeitpunkt die 50 %-Überlebenswahrscheinlichkeit erreicht wird ("Mediane Überlebenszeit").

4. Singuläre Sterbefälle erzeugen eine relative Sprunghöhe $1/n(t)$, wenn n(t) die Anzahl der bis zum Zeitpunkt t noch beoabachteten Patienten bezeichnet. (Siehe dazu insbesonder das Ende der Kurve!)

5. Zensierungen erzeugen keine Änderung der Überlebenskurven. Sie sind durch Marker gekennzeichnet.

6. Die Kurve wird mit zunehmender Beobachtungszeit immer unsicherer. Sie stellt nur eine "Schätzung" der Überlebenswahrscheinlichkeiten dar.

7. Der Standardfehler dieser Schätzungen wird in der Tabelle mit ausgedruckt (aber nicht in der Graphik dargestellt).

2.3. Statistische Vergleiche von Überlebenskurven

Häufig sollen zwei oder mehr Überlebenskurven miteinander verglichen werden, die aus den Beobachtungen von Patientengruppen mit unterschiedlichen Therapien oder mit unterschiedlichen Risikofaktoren gebildet werden.
Beispiel:

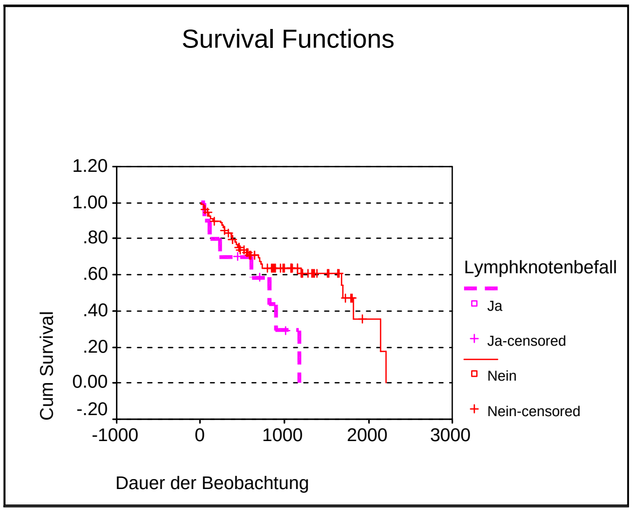

Abbildung 10

Man möchte wissen, ob die gefundenen Unterschiede der Kurven allein durch Zufall erklärt werden können oder auf unterschiedliche "Grundgesamtheiten" zurückgeführt werden müssen.

Ansatz:

In (wieder sehr kleinen) Intervallen bildet man Kreuztabellen der Form

Ereignis:	Verstorben	Zensiert	Summe:
Gruppe 1	d_1	w_1	n_1
Gruppe 2	d_2	w_2	n_2
Summe:	$d_1 + d_2$	$w_1 + w_2$	$n = n_1 + n_2$

Wenn dann die Sterbewahrscheinlichkeiten für diese Intervall in beiden Gruppen gleich wären, müssten sie so sein wie in beiden zusammen, nämlich $p = (d_1 + d_2)/(n_1 + n_2)$. In der ersten Gruppe würde man dann $n_1 \times p$ erwarten und in der zweiten Gruppe $n_2 \times p$. Beobachtet wurden aber d_1 bzw. d_2 Sterbefälle.

Man bildet nun für die erste Gruppe über alle Intervalle: die Summe aller *beobachteten* Sterbefälle: D_1, und die Summe aller *erwarteten* Sterbefälle E_1. Analog berechnet man für die zweite Gruppe die Summen D_2 und E_2. Dann berechnet man die Abweichungen der beobachteten von den erwarteten Fällen nach der Formel

$$\begin{aligned} X^2 &= \sum \frac{(\text{Beobachtete Anzahl - Erwartete Anzahl})^2}{\text{Erwartete Anzahl}} \\ &= \frac{(D_1 - E_1)^2}{E_1} + \frac{(D_2 - E_2)^2}{E_2} \end{aligned}$$

Beispiel: Siehe M. Bland, S. 287/288.

Dies ist die Formel für den logrank Test. Die Testgröße ist χ^2- verteilt mit *einem* "Freiheitsgrad", *wenn die Nullhypothese richtig ist*, dass beide Überlebenskurven aus *derselben Grundgesamtheit* stammen:

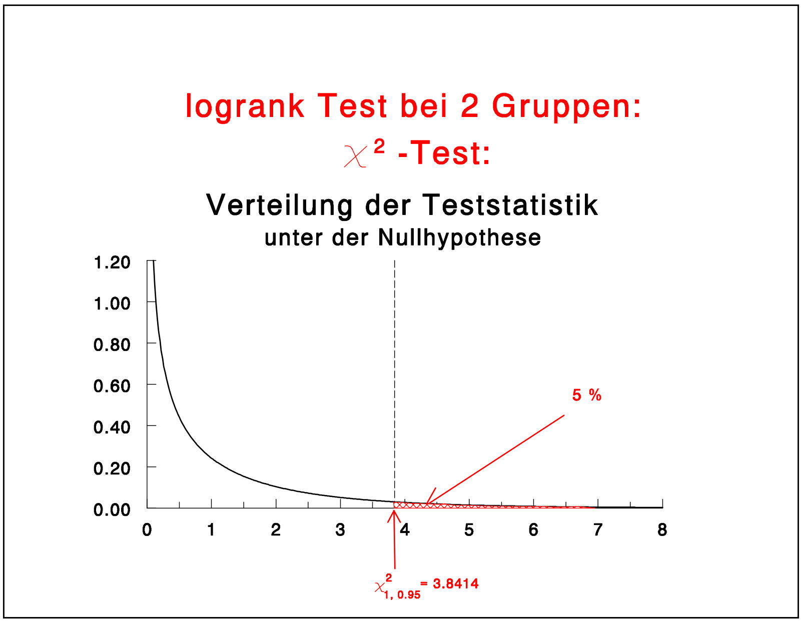

Abbildung 11

Dann ist alsodie Wahrscheinlichkeit, dass die Testgröße rein zufällig größer als 3.84 wird, nur 5%. Wegen dieser geringen Wahrscheinlichkeit wird sie dann abgelehnt und man schließt, dass die Unterschiede *mehr als zufällig* sind.

Bei noch größeren Werten von X^2 (weiter rechts auf der x-Achse) wird die Wahrscheinlichkeit noch kleiner als 5%. Man berechnet diese Wahrscheinlichkeit und nennt sie den *P-Wert der Testgröße.* Die Ablehnung der Nullhypothese erfolgt also üblicherweise, *wenn der P-Wert kleiner als 0.05 ist.*

Beispiel für die Daten der Abbildung 10:

Survival Analysis for DAUER   Dauer der Beobachtung

		Total	Number Events	Number Censored	Percent Censored
ERST63	Nein	81	32	49	60.49
ERST63	Ja	10	7	3	30.00
Overall		91	39	52	57.14

Test Statistics for Equality of Survival Distributions for ERST63

	Statistic	df	Significance
Log Rank	4.40	1	.0359

Der P-Wert ist hier also 0.0359 (< 0.05): die Unterschiede zwischen den beiden Kurven sind "signifikant".

3. Spezielle Fragestellungen

Bisher war immer als gegeben angenommen, welches der *Beginn* der Beobachtung eines Patienten ist und wie das untersuchte "*Ereignis*" definiert ist. Dies ist jedoch gerade in onkologischen Studien nicht immer so klar und sollte daher genau definiert sein. Mögliche "Kandidaten" für die einzelnen *Zeitpunkte* bei den Definitionen von "Total Survival", "Eventfree Survival", Progressfree Survival" "Time to Progress" "Diseasefree Survival" und weitere Kriterien sind:

- Beginn der Beobachtung:
  - Datum der Diagnose
  - Randomisation
  - Erste Behandlung
  - Komplette Remission, No evidence of disease
- Ende der Beobachtung/ Ereignis
  - Non-response nach 1. (2. ...) Kurs
  - Progress

- Relapse
- Neue Tumorerkrankung
- Tod durch Tox
- Tod durch Tumor
- Tod durch andere Ursachen
- Ende der Beobachtung

Die jeweils benutzte Definition sollte im Studienprotokoll zweifelsfrei beschrieben sein.