

Basis-Kurs Statistik und SPSS für Mediziner

Lösungen

SPSS-Übung – Überlebenszeitanalyse

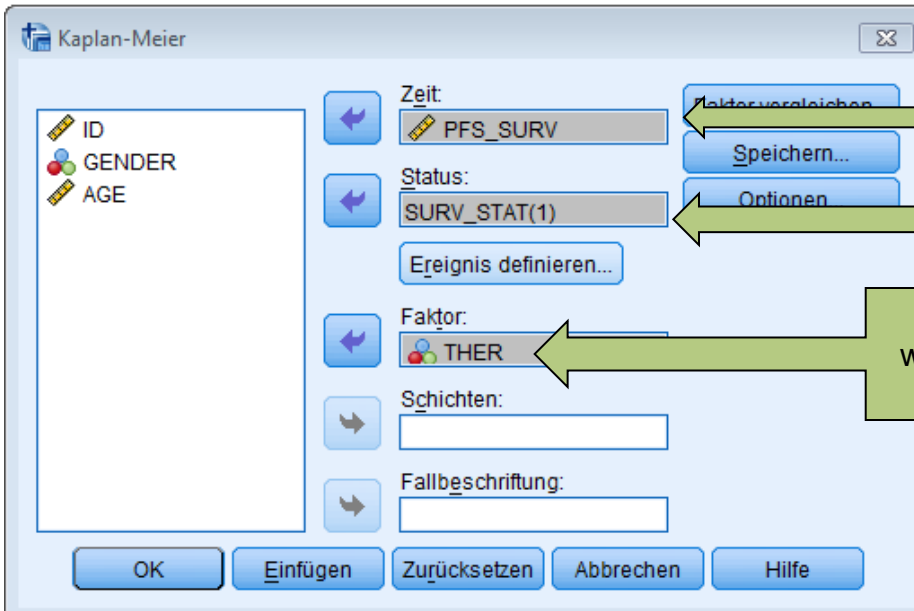
Mit Datensatz „Daten_Übung_Überlebenszeitanalyse.sav“

- 1) Zeichnen Sie die Kaplan-Meier-Kurven des progressionsfreien Überlebens (PFS_SURV mit STATUS_SURV) getrennt für beide Therapiegruppen (THER).
 - a) Wie hoch ist das mediane Überleben in beiden Gruppen?
 - b) Wie hoch ist das 5-Jahres-Überleben (60 Monate) in beiden Gruppen?

Da die Zeit des progressionsfreien Überlebens nicht für jeden Patienten bekannt ist und somit auch zensierte Beobachtungen vorliegen, müssen spezielle Methoden der Ereigniszeitanalyse verwendet werden, um in dieser Situation die vollständige Information zu berücksichtigen und unverzerrte Schätzungen zu erhalten.

Zur Schätzung und Illustration der Verteilung der Überlebenswahrscheinlichkeiten anhand der sogenannten Survivor-Funktion eignet sich die Kaplan-Meier-Methode.

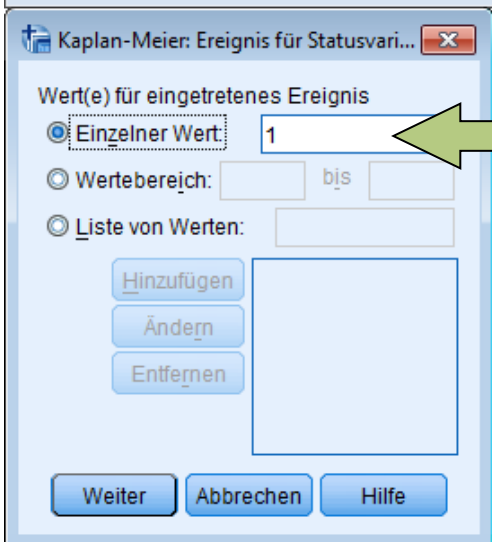
Analysieren → Überleben → Kaplan-Meier



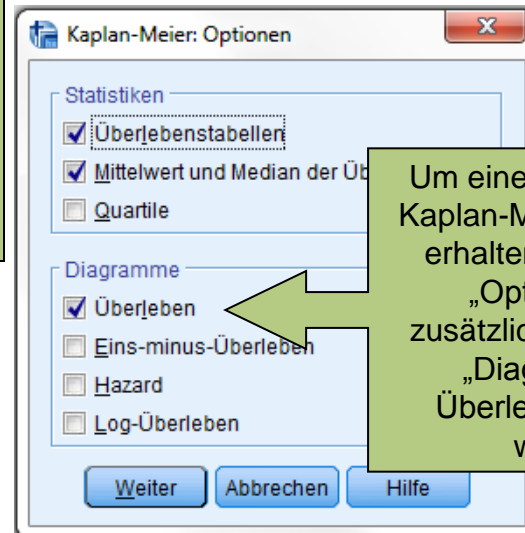
Variable mit Ereigniszeit bzw. Beobachtungszeit

Statusvariable, die anzeigt, ob es sich um eine echte Ereigniszeit oder eine zensierte Beobachtung handelt

Unter „Faktor“ kann die zu Gruppenvariable angegeben werden. Wird kein „Faktor“ spezifiziert, so wird die Survivor-Funktion für das Gesamtkollektiv geschätzt.



Unter „Ereignis definieren“ muss angegeben werden, welche Codierung ein beobachtetes Ereignis anzeigt



Um eine Abbildung der Kaplan-Meier-Kurven zu erhalten, muss unter „Optionen“ ein zusätzlicher Haken bei „Diagramme → Überleben“ gesetzt werden.

Ausgegeben werden:

Eine Übersicht der eingeschlossenen Patienten und der beobachteten Ereignisse

Zusammenfassung der Fallverarbeitung

THER	Gesamtzahl	Anzahl der Ereignisse	Zensiert	
			N	Prozent
Therapie A	62	44	18	29,0%
Therapie B	67	35	32	47,8%
Gesamt	129	79	50	38,8%

Überlebenstabelle

THER	Zeit	Status	Kumulierter Anteil Überlebender zum Zeitpunkt		Anzahl der kumulativen Ereignisse	Anzahl der verbliebenen Fälle	
			Schätzer	Standardfehler			
Therapie A	1	1,000	0	.	0	61	
	2	2,000	1	.	1	60	
	3	2,000	1	,967	,023	2	59
	4	3,000	1	,951	,028	3	58
	5	4,000	1	,934	,032	4	57
	6	4,000	0	.	.	4	56
	7	6,000	1	,918	,035	5	55
	8	6,000	0	.	.	5	54
	9	7,000	1	,904	,039	6	53

Die geschätzten Wahrscheinlichkeiten für progressionsfreies Überleben mit Standardfehlern zu den beobachteten Ereigniszeiten

Mittelwerte und Mediane für die Überlebenszeit

THER	Mittelwert ^a				Median			
	Schätzer	Standardfehler	95%-Konfidenzintervall		Schätzer	Standardfehler	95%-Konfidenzintervall	
			Untere Grenze	Obere Grenze			Untere Grenze	Obere Grenze
Therapie A	45,799	6,252	33,546	58,053	32,000	3,664	24,819	39,181
Therapie B	77,012	10,584	56,268	97,756	63,000	13,793	35,966	90,034
Gesamt	61,110	5,975	49,399	72,820	41,000	6,895	27,485	54,515

a. Die Schätzung ist auf die längste Überlebenszeit begrenzt, wenn sie zensiert ist.

Der „Mittelwert“ der Überlebenszeiten hängt stark von der Verteilung der Zensierungszeiten ab, v.a. für späte Zeitpunkte.

Daher wird häufig von der Benutzung/Präsentation der geschätzten Mittelwerte in Ereigniszeitanalysen abgeraten.

Das **mediane progressionsfreie Überleben** beschreibt den Zeitpunkt, zu dem geschätzt die Hälfte der Patienten verstorben ist (der Kaplan-Meier-Schätzer fällt unter 50%).

Es wird hier für die beiden Gruppen mit geschätztem **Standardfehler** und **95%-Konfidenzintervall** ausgegeben.

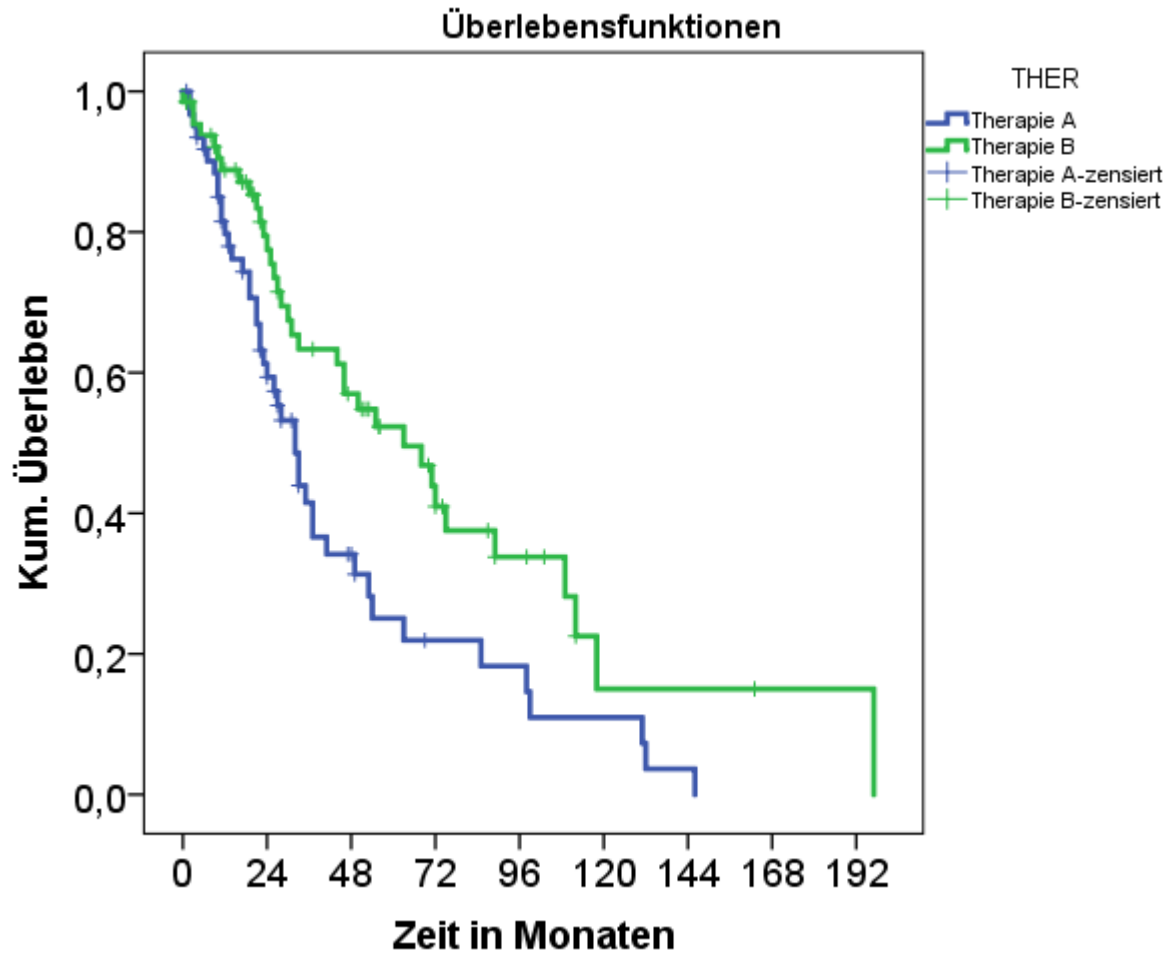
Überlebenstabelle

THER	Zeit	Status	Kumulierter Anteil Überlebender zum Zeitpunkt		Anzahl der kumulativen Ereignisse	Anzahl der verbliebenen Fälle	
			Schätzer	Standardfehler			
Therapie A	1	1,000	0	.	0	61	
	2	2,000	1	.	1	60	
	3	2,000	1	,967	,023	2	59
	4	3,000	1	,951	,028	3	58
	5	4,000	1	,934	,032	4	57
	51	49,000	1	,313	,069	33	11
	52	49,000	0	.	.	35	10
	53	53,000	1	,282	,069	36	9
	54	54,000	1	,251	,068	37	8
	55	63,000	1	,219	,066	38	7
	56	69,000	0	.	.	38	6
	57	85,000	1	,183	,065	39	5
	42	50,000	1	,548	,070	24	25
	43	51,000	0	.	.	24	24
	44	52,000	0	.	.	24	23
	45	53,000	0	.	.	24	22
	46	55,000	1	,523	,071	25	21
	47	56,000	0	.	.	25	20
	48	56,000	0	.	.	25	19
	49	63,000	1	,496	,072	26	18
	50	68,000	1	,468	,073	27	17

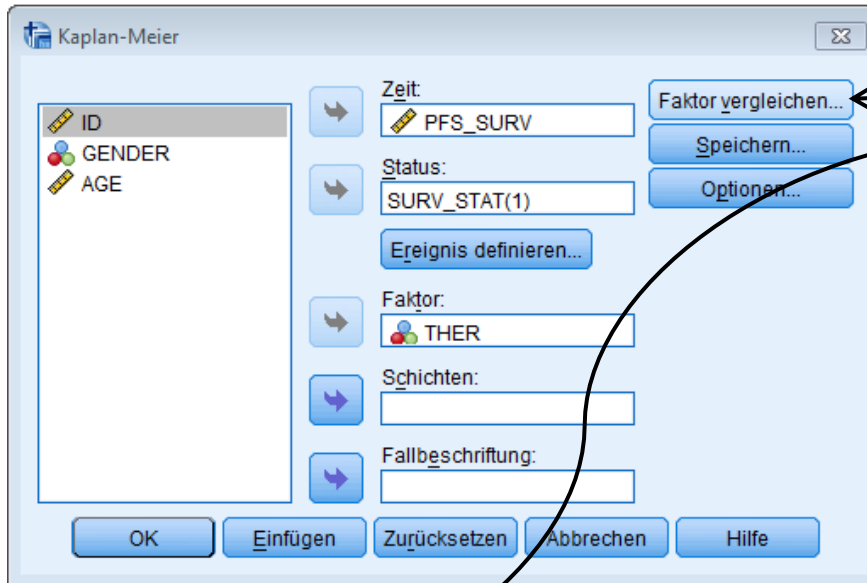
Die geschätzten 5-Jahres-Wahrscheinlichkeiten für PFS können aus der „Überlebenstabelle“ abgelesen werden. Sie entsprechen dem angegebenen Schätzer zum letzten Zeitpunkt < 60 Monate.

Die geschätzten 5-Jahres-Wahrscheinlichkeiten für PFS lauten somit:

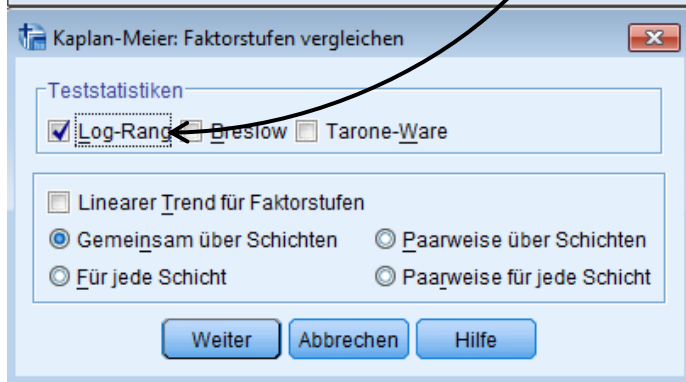
- 25,1 ± 6,8% für Gruppe A
- 52,3 ± 7,1% für Gruppe B



2) Überprüfen Sie mittels eines geeigneten Tests, ob sich die Verteilungen der Überlebenszeiten zwischen den beiden Gruppen unterscheiden. Zu welchem Schluss kommen Sie?



Unter Faktor vergleichen kann der Log-Rang zum Vergleich der Verteilung der Ereigniszeiten in unabhängigen Gruppen ausgewählt werden. Hier muss der entsprechende Haken gesetzt werden.



Da der **p-Wert** kleiner als 0,05 ist, kann die Nullhypothese („Verteilung der Ereigniszeiten ist in beiden Gruppen gleich“) auf einem Signifikanzniveau von 5% abgelehnt werden. Es gibt einen signifikanten Unterschied im PFS.

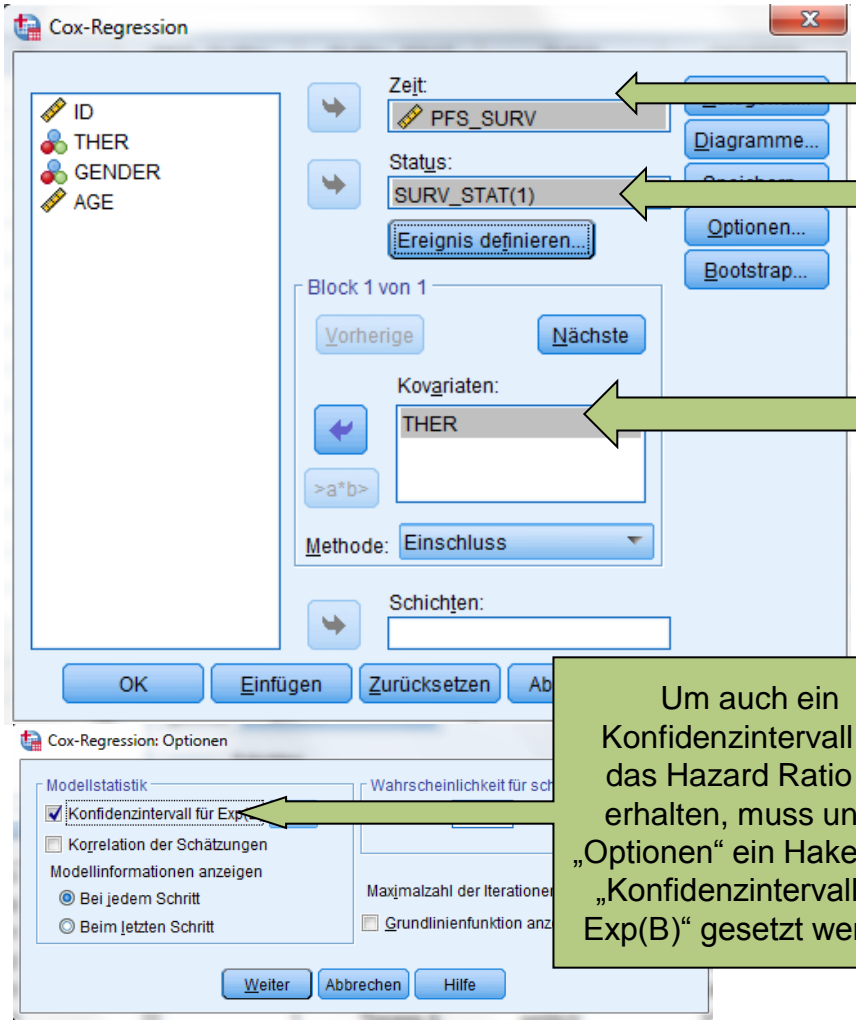
Gesamtvergleiche

	Chi-Quadrat	Freiheitsgrade	Sig.
Log Rank (Mantel-Cox)	7,329	1	,007

Test auf Gleichheit der Überlebensverteilungen für die verschiedenen Stufen von THER.

Ein Hazard Ratio (Risikoverhältnis) kann mit einem Cox-Regressionsmodell geschätzt werden:

Analisieren → Überleben → Cox-Regression



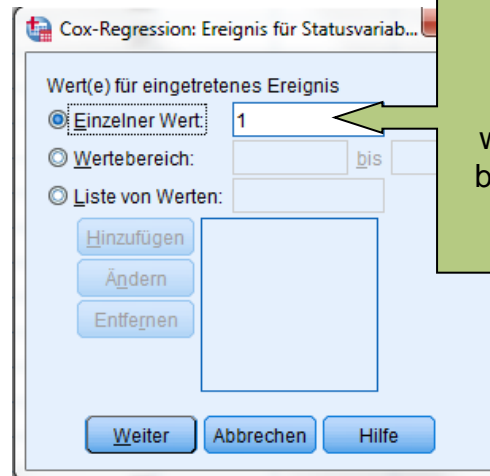
Variable mit Ereigniszeit bzw. Beobachtungszeit

Statusvariable, die anzeigt, ob es sich um eine echte Ereigniszeit oder eine zensierte Beobachtung handelt

Das interessierende Merkmal (hier Therapie – THER) als Kovariate angeben.

Um auch ein Konfidenzintervall für das Hazard Ratio zu erhalten, muss unter „Optionen“ ein Haken bei „Konfidenzintervall für Exp(B)“ gesetzt werden

Unter „Ereignis definieren“ muss angegeben werden, welche Codierung ein beobachtetes Ereignis anzeigt



Variablen in der Gleichung

	B	SE	Wald	df	Signifikanz	Exp(B)	95,0% Konfidenzinterv. für Exp (B)	
							Untere	Obere
THER	-,613	,231	7,036	1	,008	,542	,345	,852

Die relevante Information findet man in der Tabelle „Variablen in der Gleichung“.

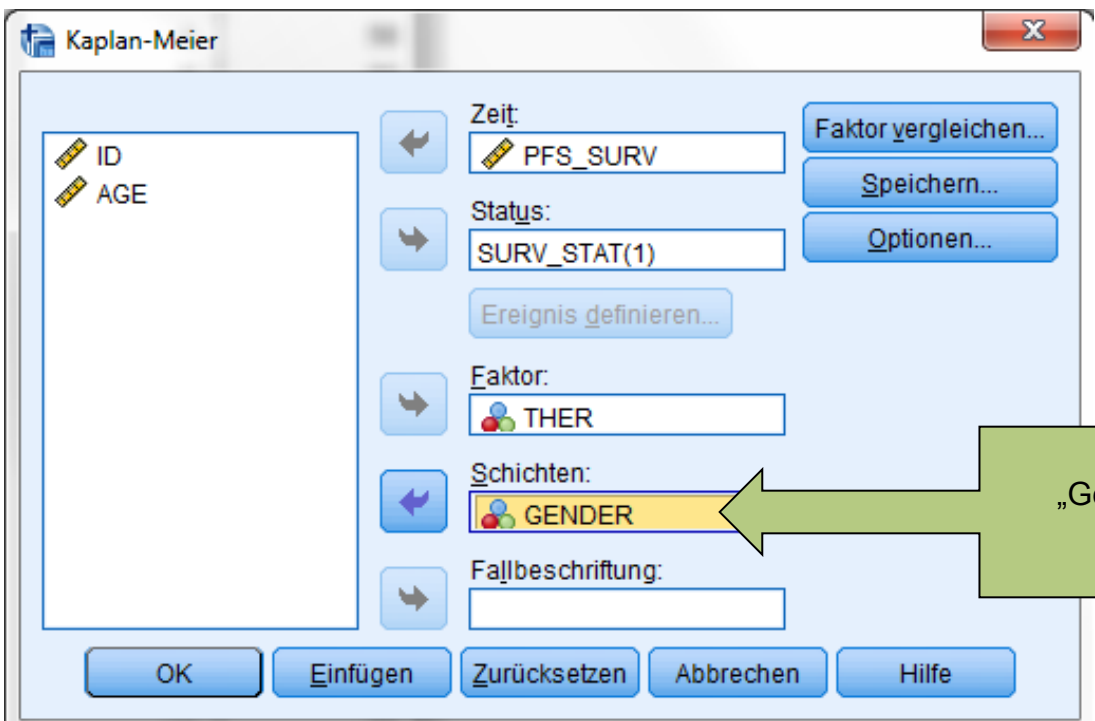
Das Hazard Ratio wird in der Spalte „Exp(B)“ angezeigt, dahinter das 95%-Konfidenzintervall.

Hier wurde als ein Hazard Ratio von 0,542 (95%-KI 0,345 bis 0,852) beobachtet. Das geschätzte Risiko in Therapiegruppe B für ein interessierendes Ereignis (Progression oder Tod) beträgt also nur 54% des Risikos in Gruppe A.

Der p-Wert der Cox-Regression ($p=0,008$) gilt als äquivalent zum p-Wert des Log-Ranks-Tests (hier: $p=0,007$).

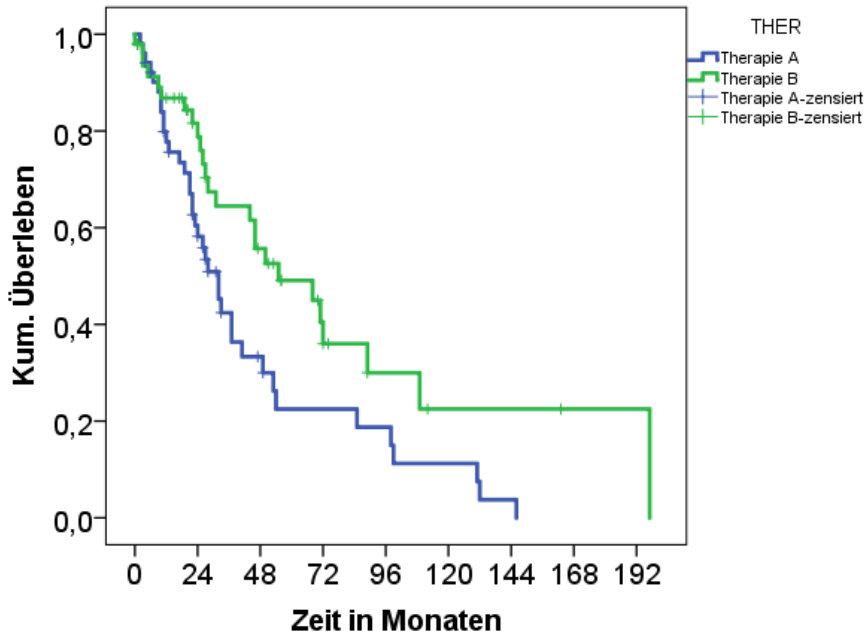
4) Zeichnen Sie die Kaplan-Meier-Kurven zum Vergleich der Therapiegruppen getrennt für Männer und Frauen und führen Sie einen entsprechenden stratifizierten Log-Rank-Test durch (Hinweis: „Schichten“).

Analysieren → Überleben → Kaplan-Meier (s. Aufgabe 1)

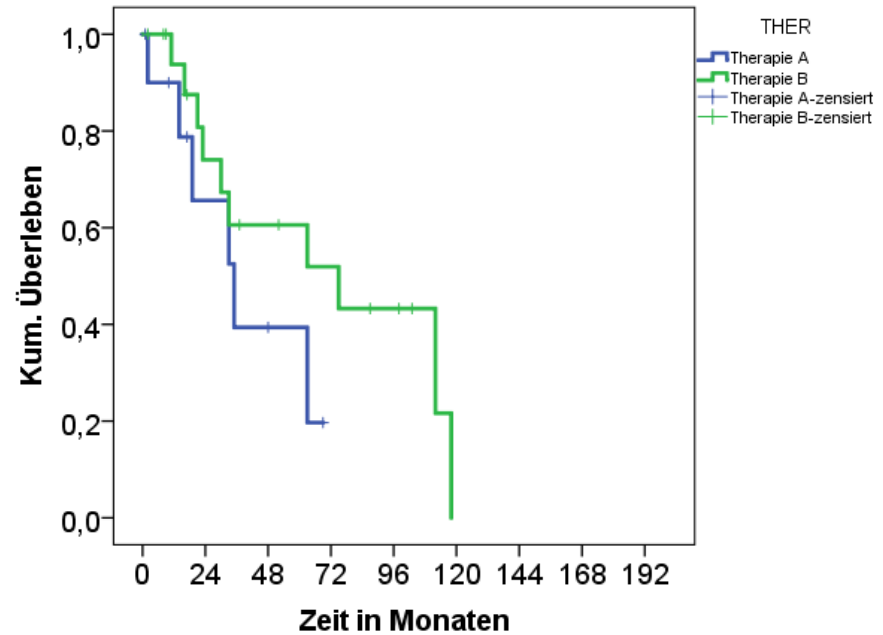


The image shows the SPSS 'Kaplan-Meier' dialog box. On the left, a list of variables includes 'ID' and 'AGE'. In the center, the 'Zeit:' field contains 'PFS_SURV', the 'Status:' field contains 'SURV_STAT(1)', and the 'Faktor:' field contains 'THER'. The 'Schichten:' field contains 'GENDER', which is highlighted in yellow. A green callout box with an arrow points to 'GENDER' and contains the text: „Geschlecht“ wird zusätzlich als Schicht (Stratum) angegeben. At the bottom, there are buttons for 'OK', 'Einfügen', 'Zurücksetzen', 'Abbrechen', and 'Hilfe'. On the right side of the dialog, there are buttons for 'Faktor vergleichen...', 'Speichern...', and 'Optionen...'.

Überlebensfunktionen
GENDER = männlich



Überlebensfunktionen
GENDER = weiblich



Gesamtvergleiche^a

	Chi-Quadrat	Freiheitsgrade	Sig.
Log Rank (Mantel-Cox)	7,387	1	,007

Test auf Gleichheit der Überlebensverteilungen für die verschiedenen Stufen von THER.

a. Korrigiert für GENDER.

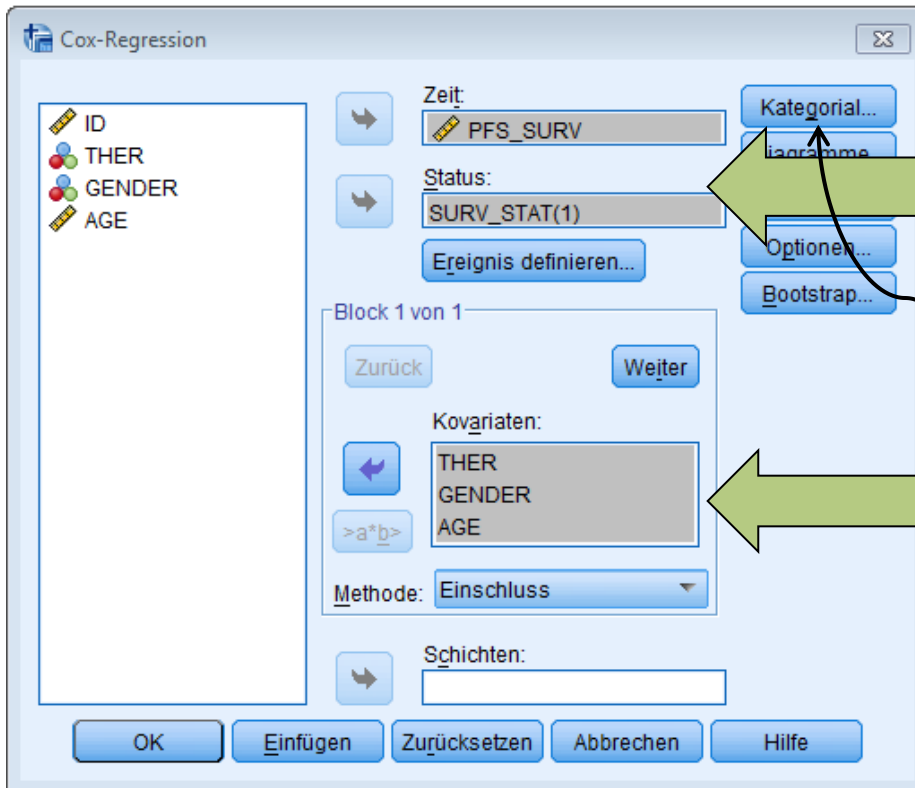
Die Kaplan-Meier-Kurven werden getrennt für die beiden Strata (männlich, weiblich) gezeichnet.

Der stratifizierte Log-Rank-Tests ist korrigiert (adjustiert) für Geschlecht. In diesem Fall ist der adjustierte p-Wert sehr ähnlich zum marginalen p-Wert ohne Adjustierung ($p_{unadj.} = 0,0068$; $p_{adj.} = 0,0065$).

Eine Stratifizierung stellt eine einfache und anschauliche Form einer Adjustierung für eine kategoriale Störgröße dar.

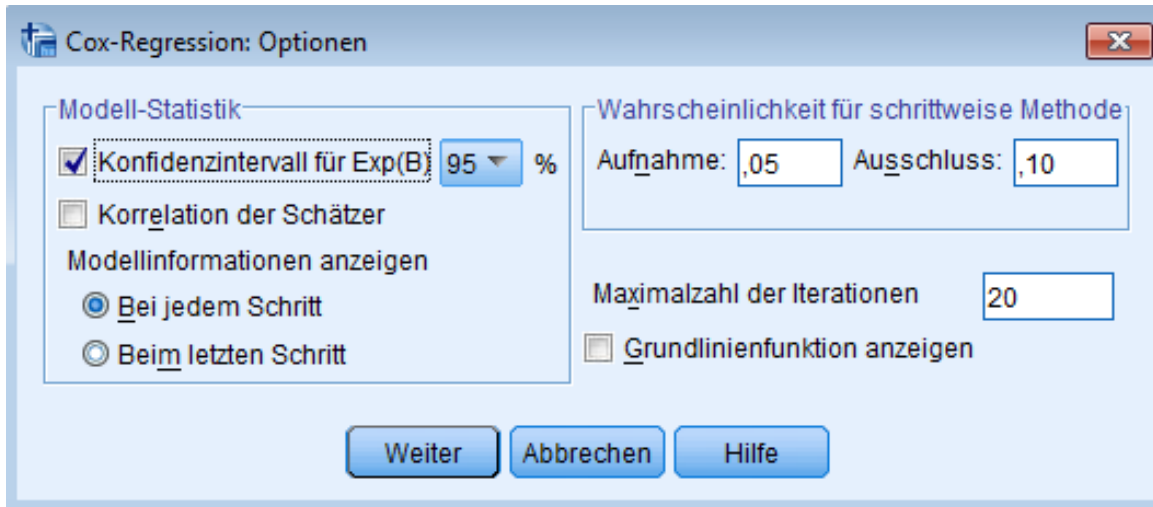
- 4) Berechnen Sie ein Regressionsmodell mit der Zielgröße progressionsfreies Überleben und den Einflussgrößen Therapie, Geschlecht und Alter und interpretieren Sie die Ergebnisse:
- Welchen Einfluss hat die Therapie?
 - Welchen Einfluss hat das Geschlecht?

Analysieren → Überleben → Cox-Regression



Zeit, Status und die Kodierung der Ereigniszeiten werden wie beim Kaplan-Meier-Schätzer festgelegt (Folie 3).

Hier können simultan mehrere (auch quantitative) Einflussgrößen berücksichtigt werden. Qualitative (kategoriale) Einflussgrößen mit mehr als zwei möglichen Ausprägungen müssen als solche gekennzeichnet werden.



Unter Optionen kann ausgewählt werden, dass auch eine 95%-Konfidenzintervall für das Hazard Ratio (=Exp(B)) geschätzt werden soll.

Variablen in der Gleichung

	B	SE	Wald	df	Signifikanz	Exp(B)	95,0% Konfidenzinterv. für Exp (B)	
							Untere	Obere
THER	-,515	,241	4,565	1	,033	,597	,372	,958
GENDER	-,050	,291	,029	1	,864	,952	,538	1,682
AGE	,018	,008	5,143	1	,023	1,018	1,002	1,034

Berücksichtigt man den gemeinsamen Einfluss der Variablen Therapie, Geschlecht und Alter, so zeigt sich:

a. Das **Hazard Ratio** zwischen der beiden Therapiegruppen beträgt 0,597 (**95%-Konfidenzintervall**: 0,372 bis 0,958). D.h. es zeigte sich ein Risikoreduktion für Therapie B (THER=1) im Vergleich zu Therapie A (THER=0) von ca. 40%. Es gibt auch nach Bereinigung möglicher Alters- und Geschlechtsunterschiede zwischen den Gruppen einen signifikanten Unterschied zwischen den Therapiegruppen: **p-Wert**=0,033 (↔ die 1 ist nicht im Konfidenzintervall enthalten).

b. Für Frauen (GENDER=1) wurde mit einem **Hazard Ratio** von 0,952 ein etwas geringeres Risiko beobachtet als für Männer (GENDER=0). Der beobachtete Unterschied ist jedoch sehr gering und nicht statistisch signifikant (**p**=0,864, **95%-KI** von 0,538 bis 1,682).

c. Für einen Altersunterschied von 10 Jahren ergibt sich ein geschätztes **Hazard Ratio** von

$$HR_{10} = \exp(10 \cdot B) = \exp(10 \cdot 0,018) = 1,197$$

also eine Risikoerhöhung von ca. 20%.