

UNIVERSITÄT LEIPZIG

Medizinische Fakultät

QSB1
Epidemiologie, Medizinische Biometrie und Medizinische Informatik

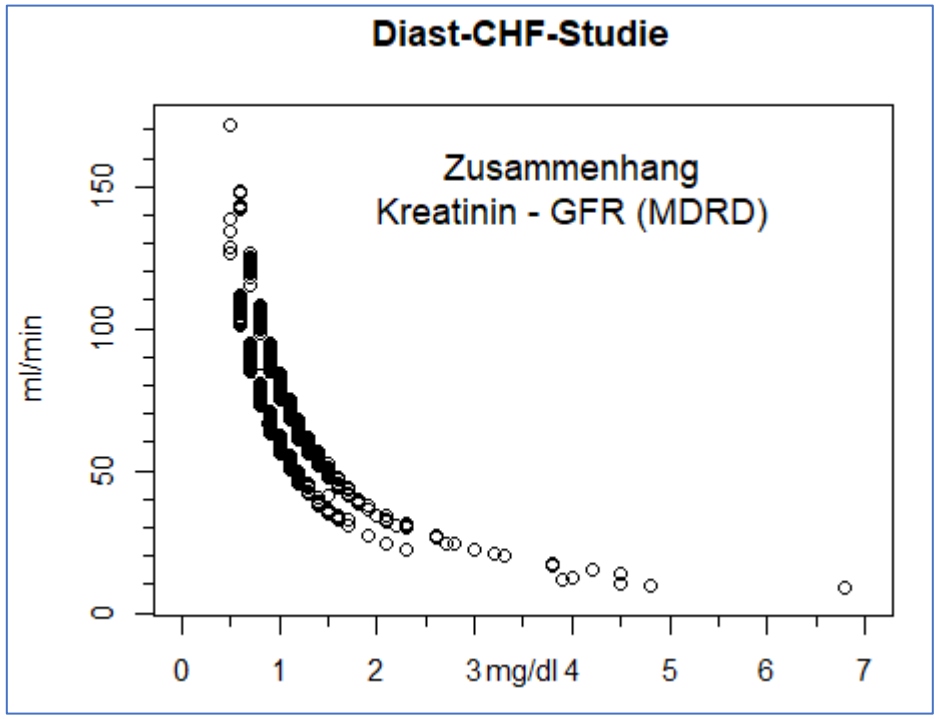
Lösungen

der Aufgaben in den Vorlesungen 7 - 13

imise.

V 7, Folie 8 Beispiel 7.2: Diast-CHF-Studie

- Zusammenhang Kreatinin - GFR



Richtung:
negativ,

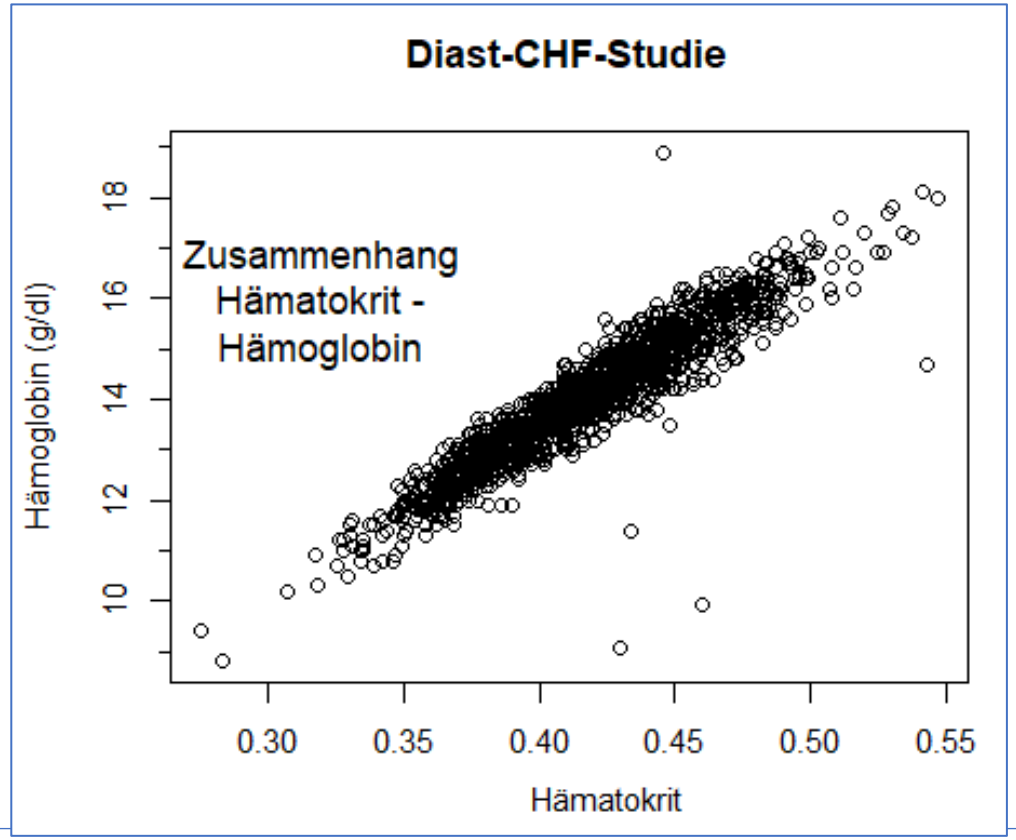
Stärke:
ziemlich starker Zusammenhang,

aber nichtlinear!

Anmerkung:
Dieser Zusammenhang entsteht durch die MDRD-Formel.

V 7, Folie 9 Beispiel 7.3: Diast-CHF-Studie

- Zusammenhang Hämatokrit - Hämoglobin

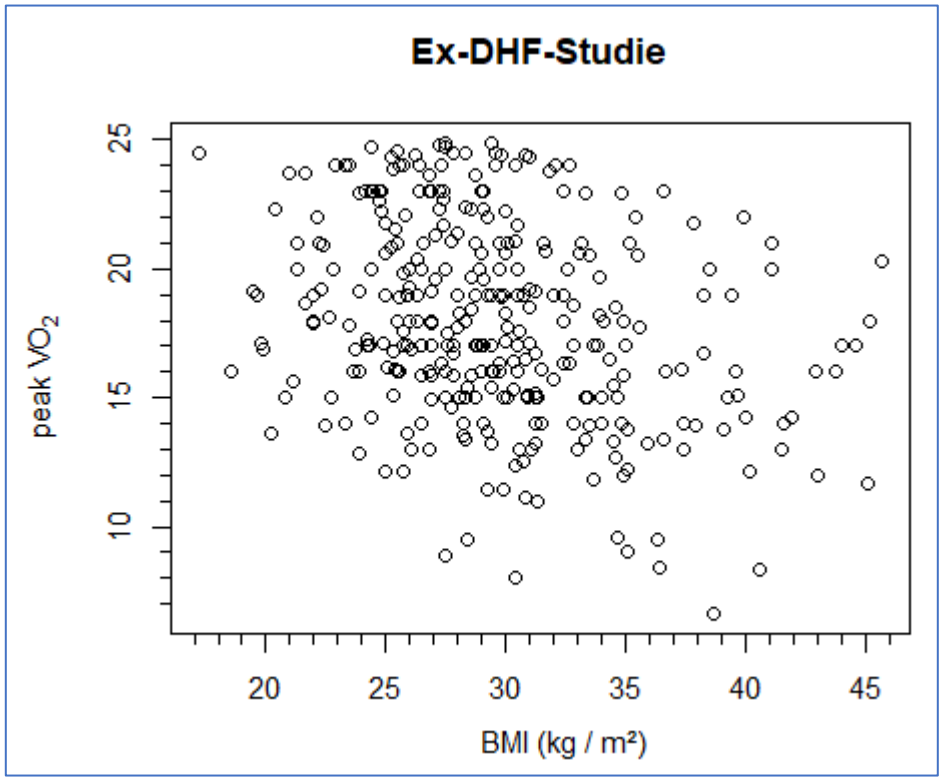


Richtung:
positiv,

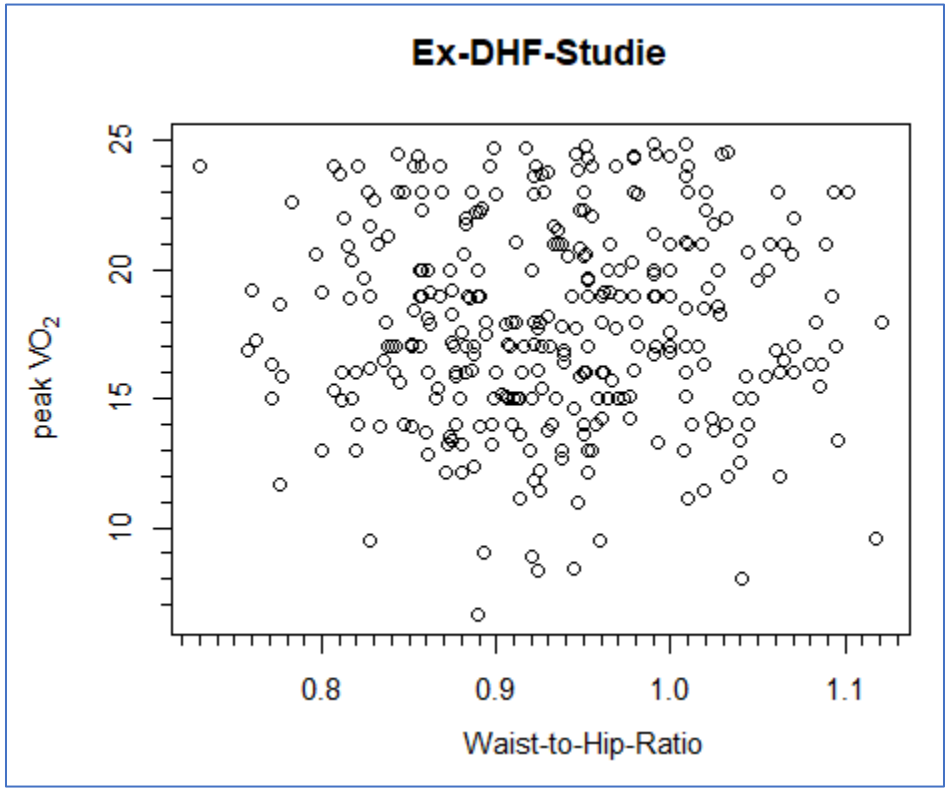
Stärke:
(sehr) starker Zusammenhang

Anmerkung:
Beide Merkmale hängen zusammen über die Erythrozyten.

V 7, Folie 10 Beispiel 7.4: Ex-DHF-Studie



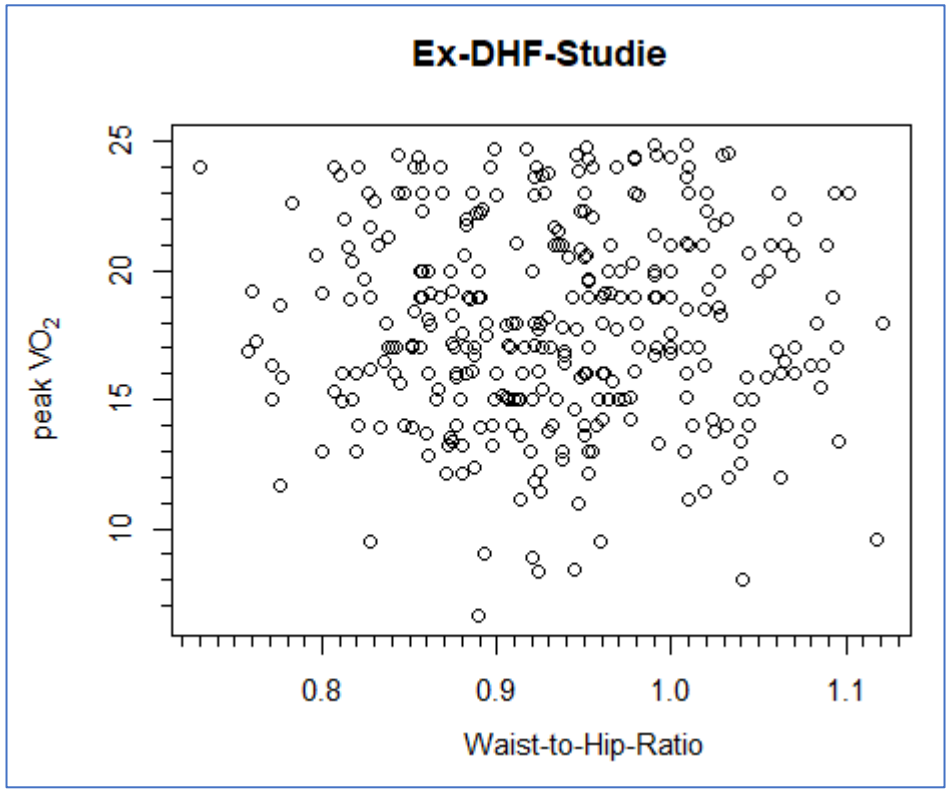
schwacher, negativer Zusammenhang



praktisch kein Zusammenhang

V 7, Folie 20 Beispiel I

- B 7.4 Quotient Taille- zu Hüft-Umfang – peak VO_2

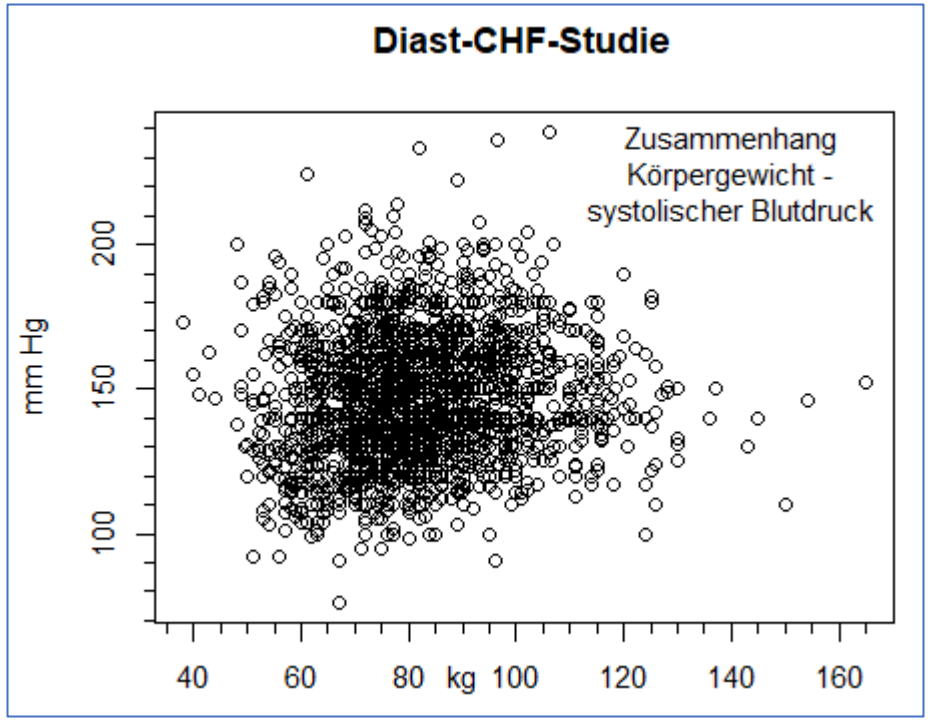


$r = -0{,}002$

keine Korrelation

V 7, Folie 21 Beispiel II

- B 7.1 Diast-CHF: Körpergewicht – systolischer Blutdruck



$r = 0{,}11$

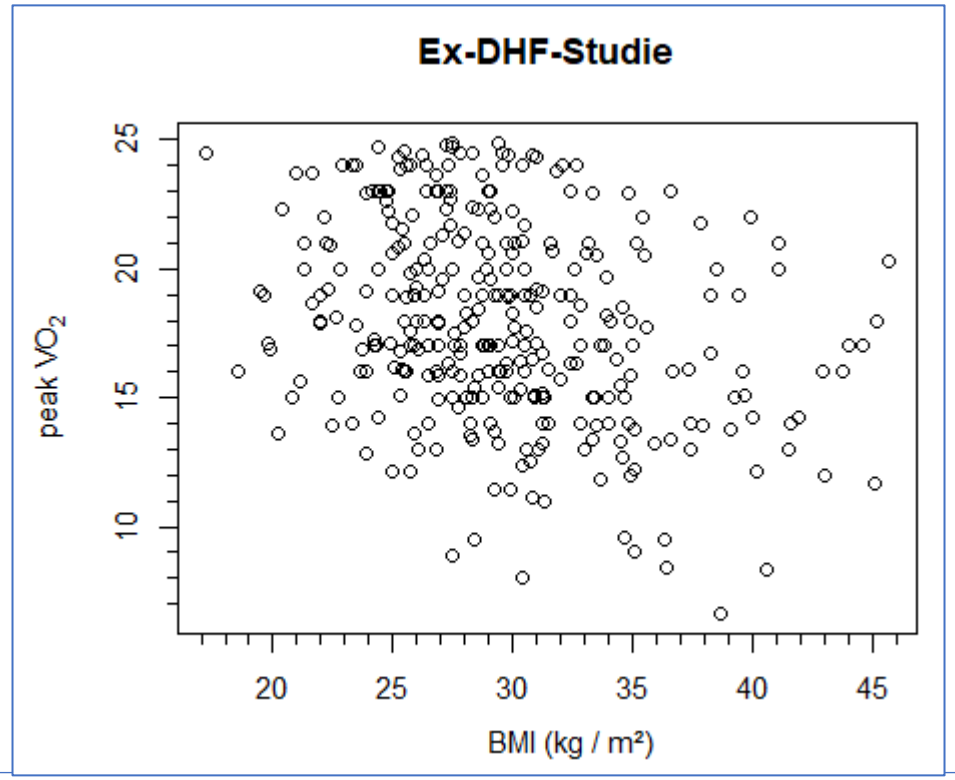
(sehr) schwache Korrelation

CAVE!
Verzerrungen ggf. durch
Geschlechtsunterschiede möglich



V 7, Folie 22 Beispiel III

- Ex-DHF-Studie: Probanden mit BMI > 20 kg/m²: BMI – peak VO_2



schwacher negativer Zusammenhang

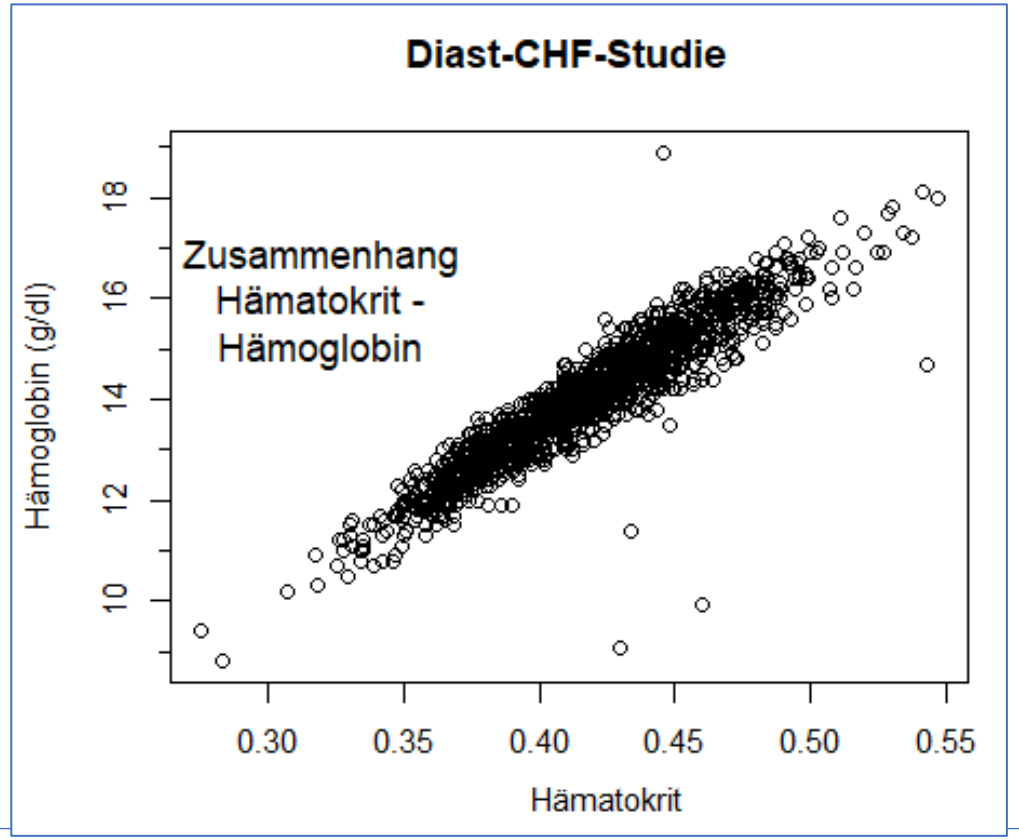
$r = -0{,}30$



Gonick & Smith: Statistik in Cartoons

V 7, Folie 23 Beispiel IV

– Diast-CHF-Studie: Hämatokrit – Hämoglobin

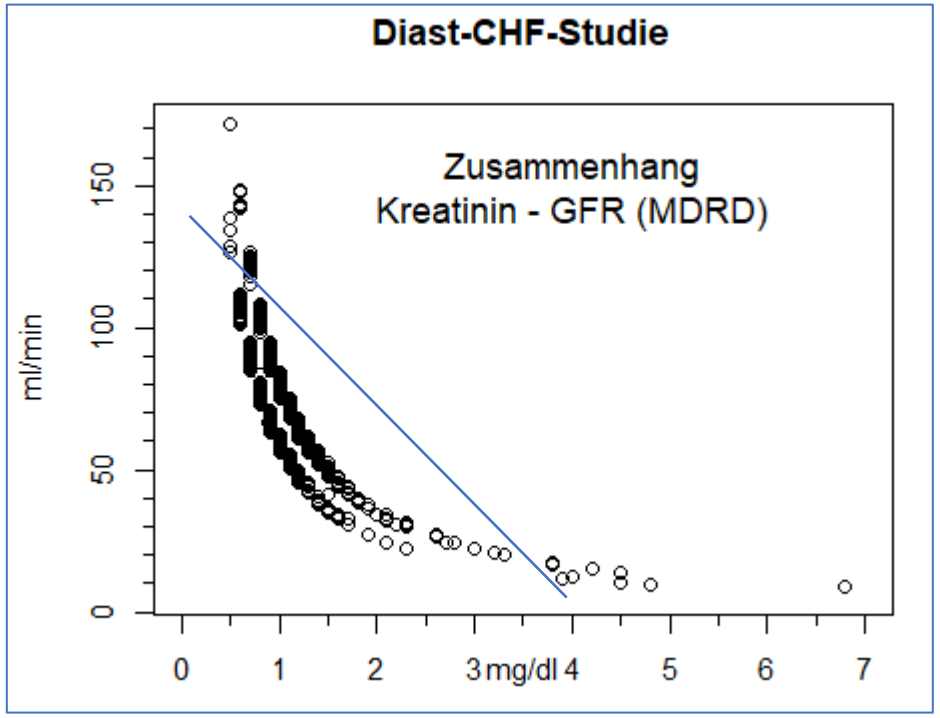


sehr starke positive Korrelation

r = 0,93

V 7, Folie 24 Beispiel V

- Diast-CHF-Studie: Kreatinin – berechnete GFR (MDRD)



starker negativer Zusammenhang

r = -0,65



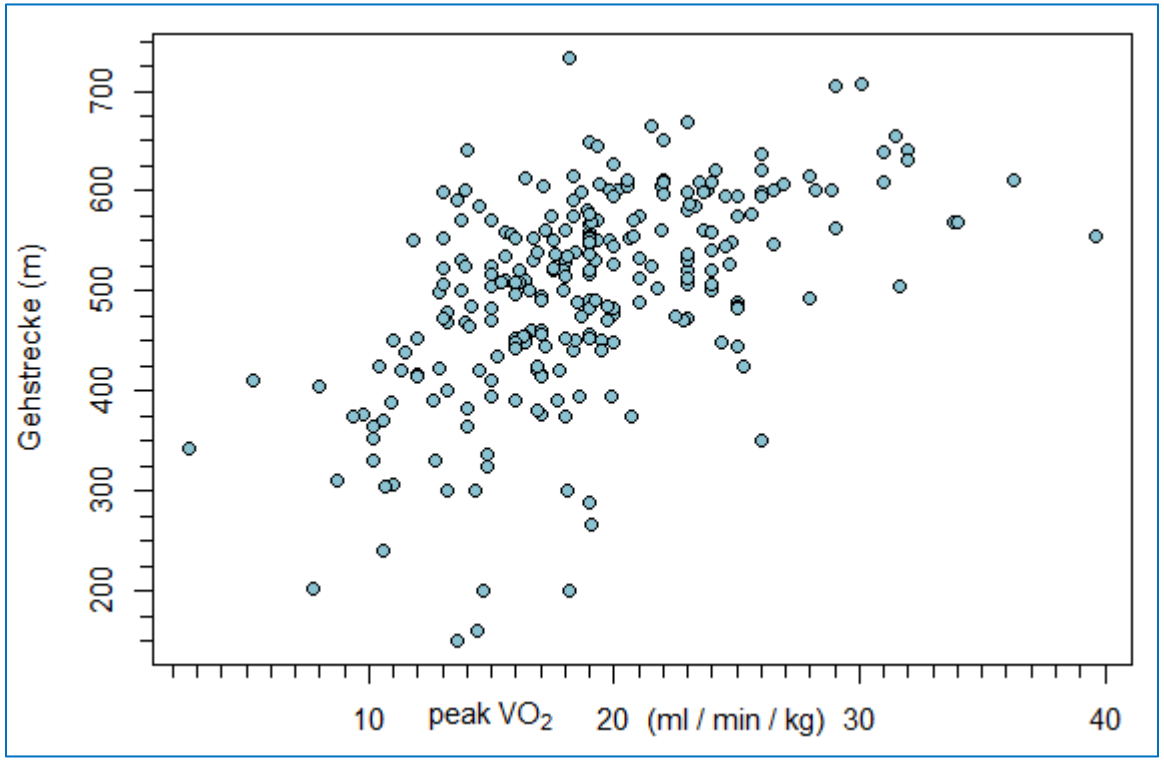
CAVE:

r misst nur den linearen Zusammenhang;

dieser ist hier aber nichtlinear!

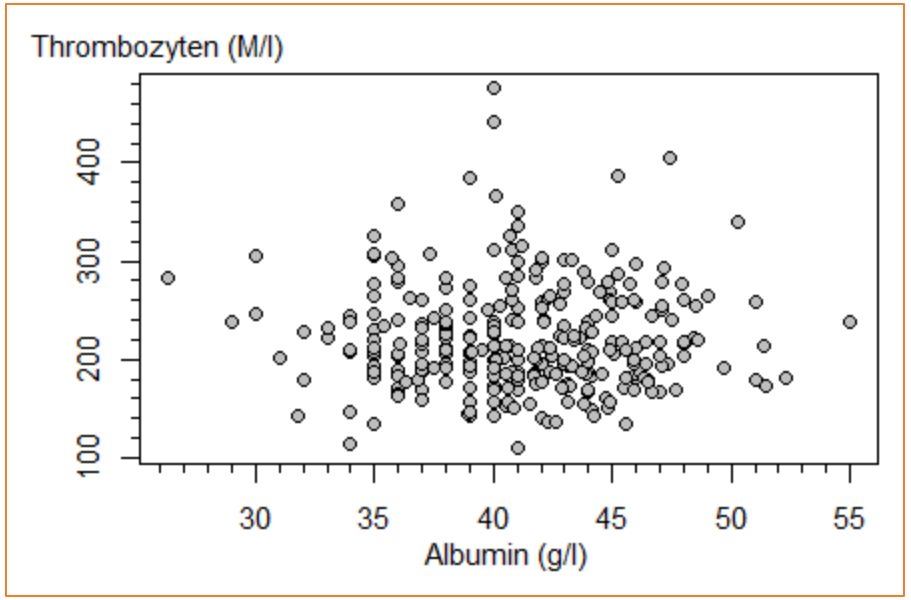
V 7, Folie 27 Übungsaufgabe 1: Erwartungshorizont

- Wie groß vermuten Sie R (Pearson)!
  - R = 0,53
  - moderater positiver Zusammenhang
- Wie interpretieren Sie das?
  - Patienten mit hohen peak-VO_2-Werten nach 12 Monaten haben im Mittel mehr Kapazität für längere 6-Minuten-Gehstrecken.



V 7, Folie 29 Übungsaufgabe 2: Erwartungshorizont

- Zwischen Albumin und Thrombozyten gibt es statistisch kaum einen Zusammenhang:
- R = - 0.01



- D. h., man könnte sagen, Albumin und Thrombozyten sind statistisch (nahezu) unabhängig.
- Doch das schließt einen möglichen biologischen Zusammenhang nicht generell aus.

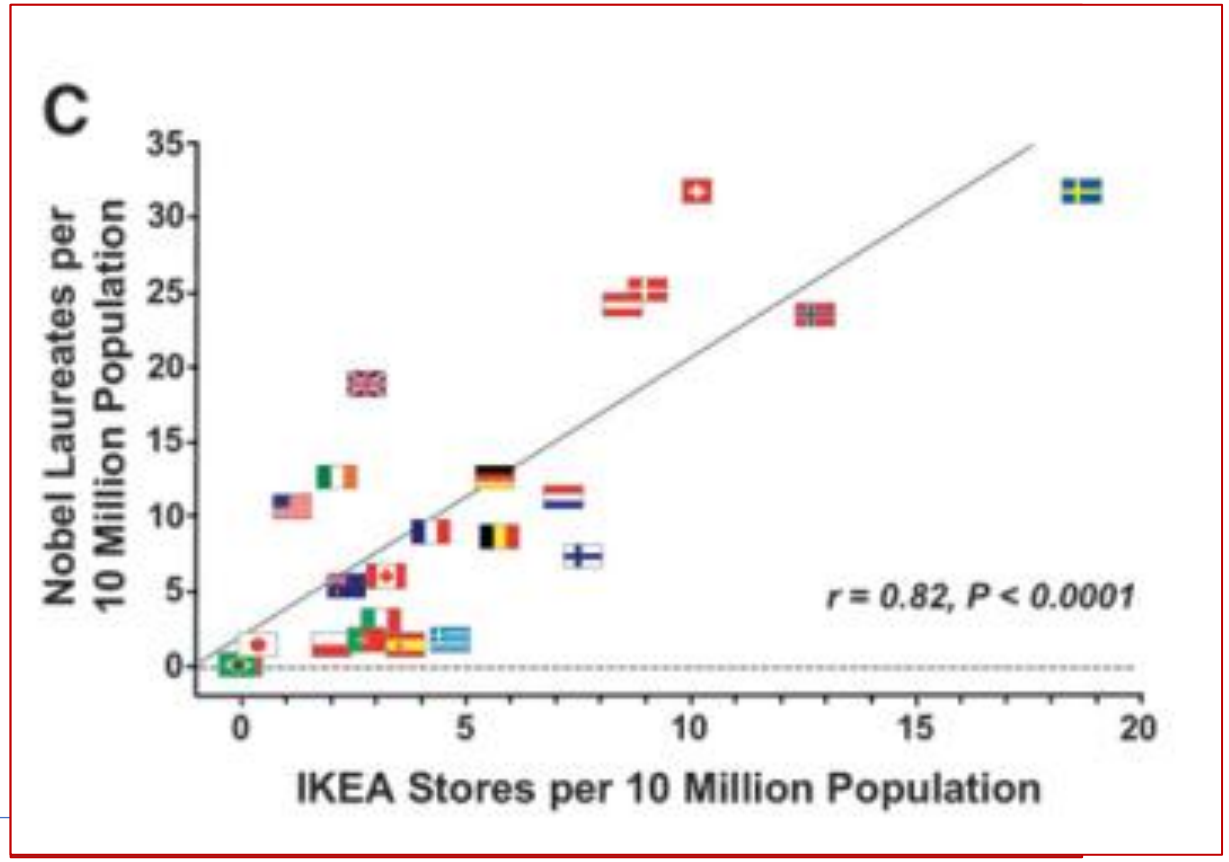
V7, Folie 50 Keine Zusammenhänge konstruieren



Mensch neigt dazu, Beobachtungen kausal zu interpretieren!

Die einfache Antwort auf diese Frage lautet, dass man, wenn man sich auf seine Intuition verlässt, häufig irrt, indem man ein Zufallsereignis als ein systematisches fehlinterpretiert.

Daniel Kahnemann. Schnelles Denken, langsames Denken.
Dt. Ausgabe: Siedler Verlag München 2011, S. 149



Messerli FH, NEJM 2012;367:1562-4

Maurage P et al. J Nutr. 2013;143(6):931-3

V 9, Folien 23f Hypothesentests: Antworten

Kreuzen Sie an bzw. beantworten Sie die Fragen!

1. Wählt man als Signifikanzniveau $\alpha = 1\%$ statt $\alpha = 5\%$, wird man im Mittel
   ○ *häufiger* ⊙ *seltener* H_0 ablehnen, wenn kein Effekt vorliegt.

2. Warum sind beim Hypothesentest Formulierungen wie
   hoch signifikant oder *borderline signifikant* unsinnig?

Das Ergebnis eines Hypothesentests ist eine (binäre) Entscheidung:
H_0 verwerfen oder nicht.

Es gibt weder ein Dazwischen noch eine quantitative Abstufung.

V 9, Folien 23f Hypothesentests: Antworten

3. Setzen Sie den Satz fort: *Die Power ist die Wahrscheinlichkeit, dass die Nullhypothese abgelehnt wird, wenn die Alternativhypothese wahr ist* (also ein Effekt vorliegt).

4. Ist $p > \alpha$ ein Beweis für die Richtigkeit der Nullhypothese?

Keinesfalls. Die Hypothesen sind nicht symmetrisch.

Erzielt man beim Hypothesentest keine Signifikanz, kann das

- an einem – gegenüber der Streuung – zu kleinen Effekt liegen oder
- daran, dass die Fallzahl für den Nachweis (eines klinisch relevanten) Effekts einfach zu klein ist.

V 9, Folie 44 Parallelen

	Hypothesen (*die Wahrheit*)	
Test	H_0	H_A
H_0 beibehalten	☺	Fehler II. Art
H_0 _ ablehnen	Fehler I. Art	☺

	Krankheitsstatus	
Test	gesund	erkrankt
Test negativ	☺	falsch negativ
Test positiv	falsch positiv	☺

Basiswahrscheinlichkeit für einen Effekt ← Prävalenz

V 10, Folie 27 Beispiel 10.3 II

Was vermuten Sie nach dem ersten Blick?

- 5 verbessert (- bedeutet Erhöhung)
- 3 verschlechtert
- 2 mit sehr starker Verbesserung

Spieler	Abfluggeschwindigkeit des Balles [km/h] vor dem Training	Abfluggeschwindigkeit des Balles [km/h] nach dem Training	Differenz
1	41	49	-8
2	44	41	3
3	48	56	-8
4	55	51	4
5	57	72	-15
6	58	52	6
7	60	66	-6
8	68	80	-12

V 11, Folie 10 Schaden durch Fehler beim Testen

Welcher Fehler (I. oder II. Art) hat schlimmere Auswirkungen?

Eine neue Substanz zur Krebsbekämpfung wird in toxikologischen Tests auf Nebenwirkungen untersucht (H_0: keine NW).

- Fehler II. Art (keine Nebenwirkung zu entdecken) wiegt schwerer als falsche Deklaration einer (unwichtigen) Nebenwirkung.

Die Behandlung mit einer neuen Substanz zur Krebsbekämpfung wird mit Standardbehandlung verglichen (H_0: kein Unterschied im rezidivfreien Überleben).

- Fehler I. Art: → Patienten würden anhand eines falsch signifikanten Ergebnisses vergeblich mit einer unwirksamen Therapie behandelt.
- Fehler II. Art: → eine potenziell wirksame Substanz wird nicht weiter beforscht.

V 11, Folie 24 Testentscheidung

- Quantil: s. Skriptenheft, Tab. 11.6, S. 181:
- $\chi^2_{1;0,95} = 3{,}84$

Tabelle 11.6: *Quantile der χ^2-Verteilung*

FG	5 %	1 %	0,1 %	FG	5
1	3,84	6,63	10,83	51	68,
2	5,99	9,21	13,82	52	69,
3	7,81	11,34	16,27	53	70,
4	9,49	13,28	18,47	54	72,

Es gilt: $X^2 = 3{,}94 > 3{,}84 = \chi^2_{1;0,95}$

- Also wird H_0 abgelehnt.

Interpretation:

- Im R-ARP-Arm erreichen 8,7% mehr Patienten nach 12 Wochen Kontinenz,
- das ist signifikant zum 5%-Niveau (p = 0,047).

V 11, Folie 25 Übungsaufgabe

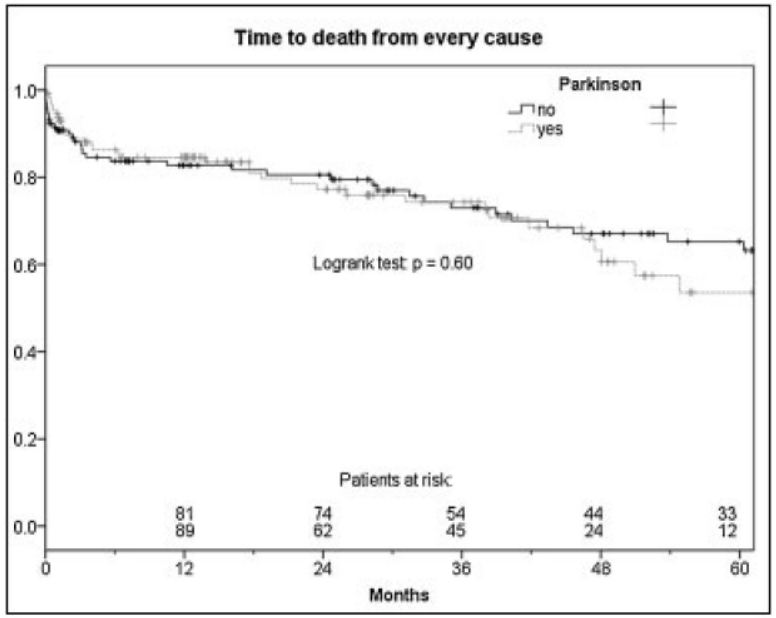
- Wie viele Freiheitsgrade hat eine X²-Statistik für eine 4 x 2-Tabelle?

- Lösung:

$z = 2, s = 4;$
also $f = (2 - 1) \cdot (4 - 1) = 3$

V 12-13, Folie 40 B 13.1 Mögliche Lösung 1

Schröter T et al. Thorac Cardiovasc Surg 2018;66:629–636



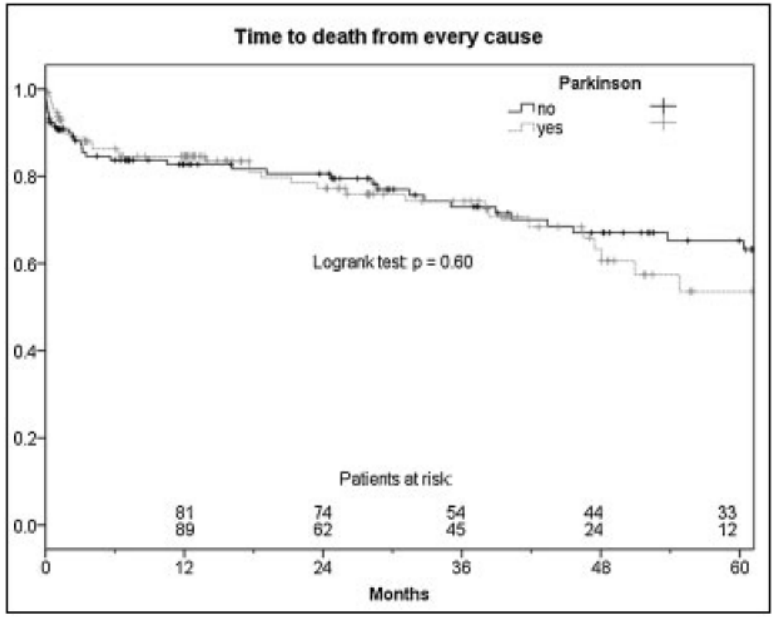
1. Kann man argumentieren: *Parkinson verschlechtert die Prognose für das 5-Jahres-Überleben deutlich.*?

Nein.

- Die Kaplan-Meier-Kurven sind bis 4 Jahre nahezu identisch. (Das korrespondiert zum p-Wert $p = 0{,}6$.)
- Der Abfall über 10% im fünften Jahr kommt durch die geringe Zahl von Patienten unter Risiko zustande. In dieser Phase sind beide Kurven bereits instabil, d. h. die Schätzwerte der Überlebenswahrscheinlichkeit sind verzerrt.

V 12-13, Folie 41 B 13.1 Lösung 2

Schröter T et al. Thorac Cardiovasc Surg 2018;66:629–636



2. Wie bewerten Sie den Satz:
   Patienten ohne und mit Parkinson haben nach Herz-OP identische Überlebenschancen.

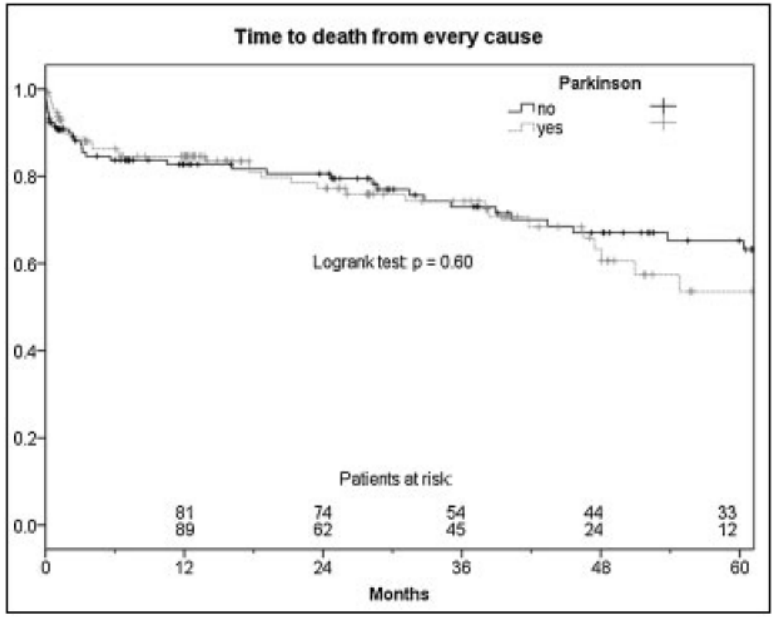
Das ist etwas zu stark formuliert, besser:

- *In der 5-Jahres-Letalität nach Herz-OP konnten wir keinen Unterschied zwischen Patienten ohne bzw. mit Parkinson nachweisen.*
- Der nicht signifikante Test ist kein Beweis dafür, dass es keinen Unterschied gibt; nur ist dieser nicht statistisch nachweisbar.

V 12-13, Folie 42 B 13.1 Lösung 3

Schröter T et al. Thorac Cardiovasc Surg 2018;66:629–636

3. Welche möglichen Verzerrungs-faktoren könnten das Ergebnis beeinflusst haben?



Typische Verzerrungsfaktoren (Confounder) fürs Überleben sind Geschlecht, Alter, Raucher, Diabetes u. a. Komorbiditäten.

– Mehrere dieser Merkmale wurden aber bereits durch das Matching balanciert (→ Strukturgleichheit)

Characteristics	Non-Parkinson (n = 130)		Parkinson (n = 130)	
Age (y), mean SD	74.8	6.6	74.9	6.6
Logistic EuroSCORE, mean SD	14.1	14.2	14.3	14.3
Body mass index (kg/m^2), mean SD	26.9	3.7	26.9	4.4
Glomerular filtration rate (Cockcroft–Gault), mean SD	66.1	23.0	66.0	24.0
Ejection fraction (%), mean SD	54	15	53	12

V 12-13, Folie 44 B 13.2 Mögliche Lösung

Welche Aussagen (Effekt, Evidenz, …) können Sie aus der Grafik ablesen?

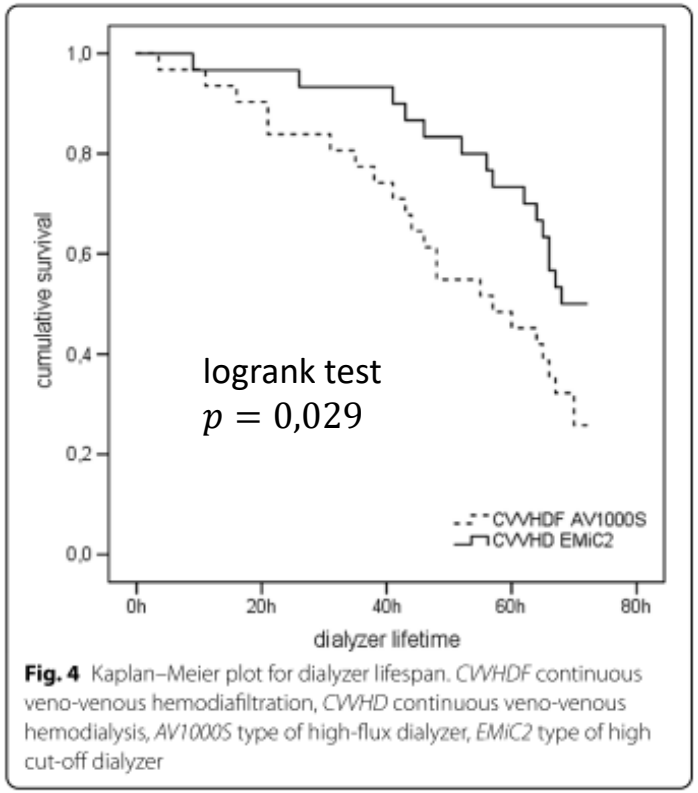


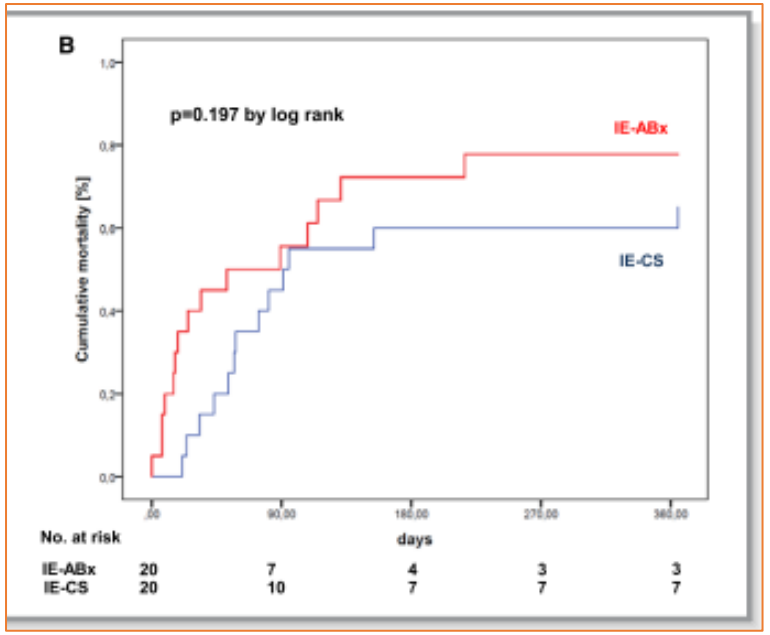
Fig. 4 Kaplan–Meier plot for dialyzer lifespan. *CVVHDF* continuous veno-venous hemodiafiltration, *CVVHD* continuous veno-venous hemodialysis, *AV1000S* type of high-flux dialyzer, *EMiC2* type of high cut-off dialyzer

- Über 10% mehr Filter des (neuen) Verfahrens CVVHD-EMC2 bleiben 3 Tage funktionsfähig als bei Kontrollen. Dieser Unterschied ist signifikant zum 5%-Niveau.
- Es lohnt sich also, mit dem neuen Filter weiter zu forschen. Anmerkung: Die Myoglobin Clearance des neuen Filters war zu allen Zeitpunkten (1, 6, 12, 24, 48 h) höher als jene des bisherigen Standards CVVHDF AV 1000S.

imise.

V 12-13, Folie 46 B 13.3 Lösung

Magner N et al. J Am Heart Assoc. 2018;7:e010027



Wie beurteilen Sie das Studienergebnis?

- Die Letalität ist in den ersten Wochen in der ABx-Gruppe deutlich höher.
- Nach etwa 4 Monaten flachen beide Kurven ab.
- Der Unterschied (75% vs. 65%) ist zum 5%-Niveau nicht signifikant.
- Es gibt also keine Evidenz für einen Vorteil des chirurgischen Verfahrens (Fallzahl!).

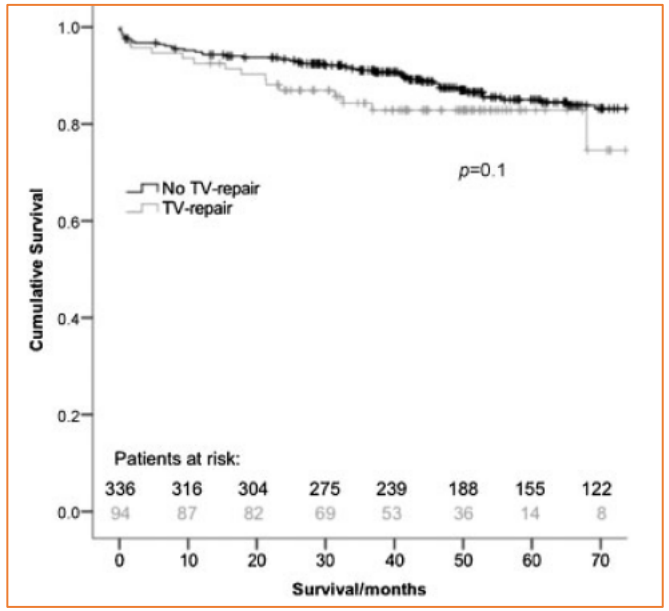
Hinweis: Kaplan-Meier-Kurven sind wegen der Ungenauigkeit der Schätzungen für solche Fallzahlen *eigentlich* nicht möglich!

B 13.4 Lösung

Pfannmüller B et al. Thorac Cardiovasc Surg 2013;61:386–391

Wie beurteilen Sie das Ergebnis der Studie?

- In 3 Jahren baut sich ein Überlebens-Unterschied von ca. 10% auf, der aber bis 5 Jahre wieder verschwindet:
- Proportional Hazards gilt nicht!
- Prinzip der Beobachtungsgleichheit ist verletzt: kürzere Nachbeobachtungszeiten in der TV-repair-Gruppe



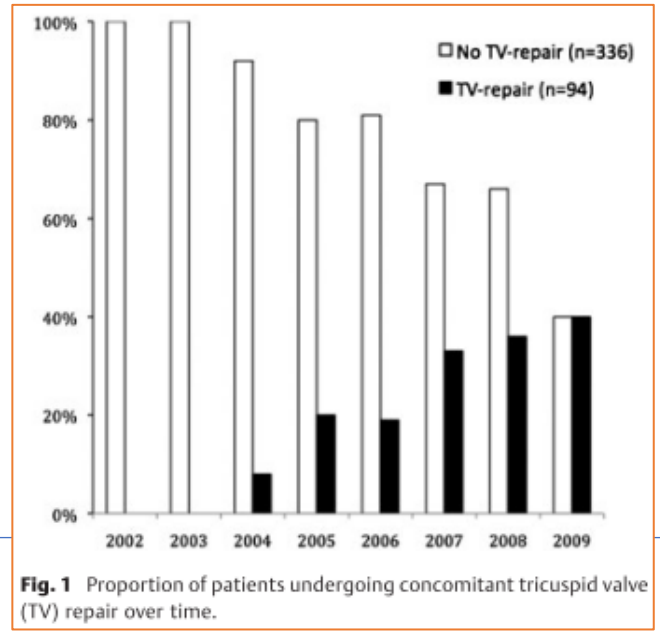


Fig. 1 Proportion of patients undergoing concomitant tricuspid valve (TV) repair over time.