

# Wie interpretiere ich medizinische Literatur?

PD Dr. med. Ueli Güller, MHS  
Universität Basel  
University of Toronto

## Wichtigkeit

- Publikationen
- Evidence Based Medicine
- Klinischer Alltag

## Journal

- Peer-reviewed vs nicht peer-reviewed
- Editor-in-Chief
- Impact factor
- Auflage

## Titel

- Uninteressant/irrelevant: **Papierkorb**
- Interessant/relevant: **=> Autoren**

## Abstract/Schlussfolgerung

- Irrelevant: **Papierkorb**
- Relevant: **=> Methodik**

## Methodik

Wichtigste Teil einer medizinischen Studie.

## Levels of Evidence

- Meta-analyse Level IA
- Randomisierte kontrollierte Studie Level IB
- Kohortenstudie Level II
- Fall-Kontrollstudie Level III
- Fallserie (case series) Level IV
- Fallbericht Level V

## Levels of Evidence

Gold Standard:

Gut geplante und durchgeführte  
prospektiv randomisierte Studie

## Levels of Evidence

Aber:

Auch prospektiv randomisierte  
Studie hat Nachteile/Limitationen!

## Prospektiv vs Retrospektiv

- Prospektiv nicht zwingend besser
- Retrospektiv nicht zwingend "low quality"

## Nota Bene

Schlecht durchgeführte prospektiv  
randomisierte Studie

<<

gut durchgeführte retrospektive Studie

## Methodik Caveats

- Statistik
- Fallzahl
- Bias

## Methodik : Statistik

„Statistics are the back bone of each and every study.“

Brent Aaron Blumenstein, PhD

## Nota Bene

„Statistics are to surgeons like a lamp post to a drunk –

more support than illumination!“

## Statistik Literatur

- Intuitive Biostatistics, H. Motulsky, Oxford University Press
- Multivariable Analysis, Mitchell H. Katz, Cambridge University Press
- Güller U et al. Interpreting Statistics in Medical Literature. JACS 2004; 198 (3): 441.
- Güller U, Oertli D: Sample Size Computations, WJS 2005; 29: 601.
- Güller U: Surgical Outcomes Research, WJS 2006;30: 255.

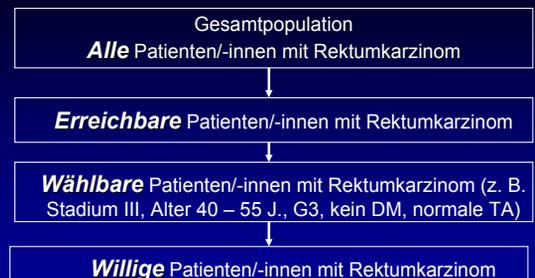
## Methodik : Fallzahl

- Fuer prospektiv randomisierte Studien
- Wurde eine Fallzahl berechnet?
- Ist berechnete Fallzahl = eingeschlossene Patienten?
- Ist die Studie underpowered?

## Methodik : Bias

- Selektionsbias
- Confounding
- Post-Randomisationsbias

## Selektionsbias



Hell K. Zentralbl Chir 2000;125 Suppl 2:196.  
Bailey KR. Control Clin Trials 1994; 15: 15.  
Kennedy WA. Int J Tech Ass Health Care 2003; 19: 8.

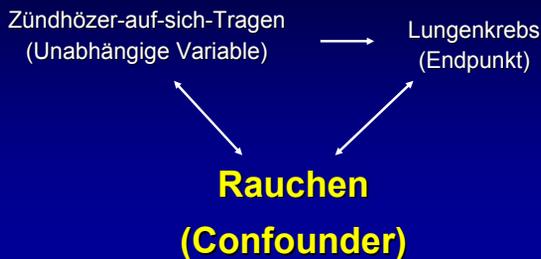
## Confounding

- **Unabhängige Variable:** z. B. Medikament, Operationstechnik, etc.
- **Abhängige Variable (Endpunkt, Outcome):** z. B. Gesamtüberleben, Rezidivrate, postoperative Komplikationen, etc.

## Definition Confounding

Situation, in welcher der Zusammenhang zwischen einer unabhängigen Variable und einem Endpunkt (abhängige Variable) auf ein drittes Merkmal zurückzuführen ist.

## Confounding



## Confounding

### Stratifizierung:

Analyse der Assoziation zwischen Zündhölzer-auf-sich-Tragen und Lungenkrebs nach Unterteilung der Patienten in Raucher vs Nichtraucher.

## Confounding

Bei mehreren Confounders:

Multivariate Analyse >> Stratifizierung

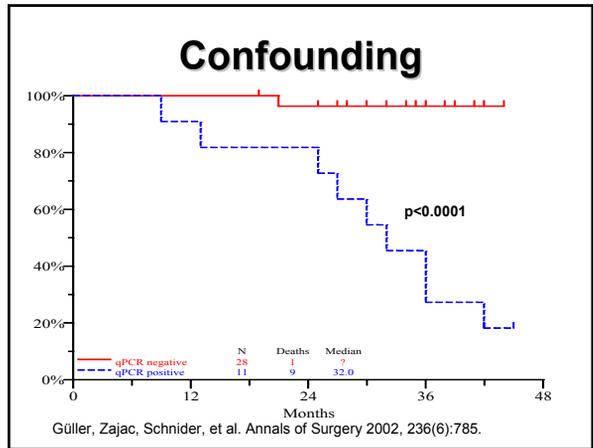
## Confounding

**Beispiel:** Prognostische Bedeutung von im Blut disseminierten Tumorzellen in Kolorektal-Karzinompatienten/-innen.

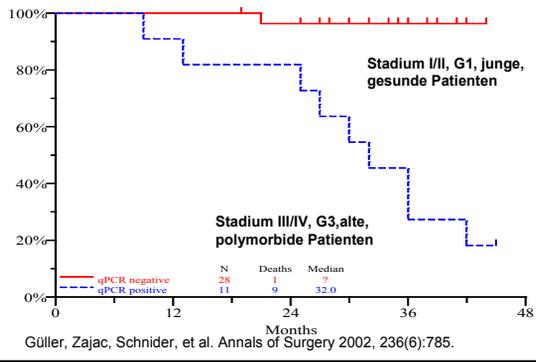
**Hypothese:** Patienten/-innen mit disseminierten Tumorzellen haben schlechteres Gesamtüberleben im Vergleich zu Patienten/-innen ohne disseminierte Tumorzellen.

# Confounding

Präsenz von disseminierten Tumorzellen → Gesamtüberleben



# Confounding



# Confounding

Präsenz von disseminierten Tumorzellen → Gesamtüberleben

Alter, Nebenerkrankungen, Tumorstadium, Grading  
(Confounders)

# Post-Randomisations Bias

Patienten/-innen mit St. n. rez. Sigmadivertikulitis  
Lap. Sigmaresektion → Offene Sigmaresektion

Outcome: Postoperative Komplikationen

# Post-Randomisations Bias

Patienten/-innen mit St. n. rez. Sigmadivertikulitis  
Lap. Sigmaresektion → Offene Sigmaresektion

5% vs 15%, p<0.001

Outcome: Postoperative Komplikationen

## Post-Randomisations Bias



Deveraux PJ. BMJ 2005; 330: 88.  
Lacy AM. Lancet 2002; 359: 2224.

## Post-Randomisations Bias

Differenz bez. Outcome wird durch unterschiedliche chirurgische Expertise verwischt (biased).

## Methodik

- Stark biased: **Papierkorb**
- Korrekt **=> Resultate**

## Resultate

**F** easible  
**I** nteresting  
**N** ovel  
**E** thical  
**R** elevant

## Resultate

- Ist die Arbeit relevant?
- Ist die Fragestellung interessant?
- Ist die Hypothese neu?
- **Nein: Papierkorb**
- **Ja: Weiterlesen**

## Resultate Caveats

- P-Werte
- Statistische Signifikanz vs klinische Relevanz
- Relative- vs absolute Risikoreduktion

## P-Wert: Nota Bene

Der p-Wert ist stark von der Fallzahl abhängig.  
Je grösser die Fallzahl, desto signifikanter der p-Wert.

## P-Wert: Nota Bene

Ein nicht signifikanter p-Wert demonstriert nicht unbedingt, dass keine Differenz besteht, sondern nur, dass die Evidenz nicht stark genug ist, eine Differenz statistisch signifikant nachzuweisen (fehlende Power).

## P-Wert: Nota Bene

Das Signifikanzniveau (0.05) darf nicht als absolute Schnittstelle zwischen wichtigen und unwichtigen Resultaten betrachtet werden.

## Resultate Caveats

Statistische Signifikanz versus klinische Relevanz.

## Statistische Signifikanz vs klinische Relevanz

Spitalaufenthalt nach laparoskopisch vs offener Appendektomie:  
3.16 Tage vs 3.20 Tage: Differenz 0.04 Tage (ca. 1 Stunde):

**Klinisch irrelevant!**

## Statistische Signifikanz vs klinische Relevanz

Spitalaufenthalt nach laparoskopisch vs offener Appendektomie:  
3.20 Tage vs 3.16 Tage: Differenz 0.04 Tage (ca. 1 Stunde): Klinisch irrelevant!

- 4'908 Patienten/-innen:  $p=0.05$
- 10'678 Patienten/-innen:  $p=0.001$
- 14'004 Patienten/-innen:  $p=0.0001$

## Statistische Signifikanz vs klinische Relevanz

Bei grosser Fallzahl können auch klinisch irrelevante Unterschiede statistisch signifikant werden.

## Statistische Signifikanz vs klinische Relevanz

Randomisierte Studie Kolonkarzinom:  
5FU versus Folfox:

- 5-Jahres Gesamtüberleben: 40% versus 60%:  
Differenz 20%:

**klinisch relevant!!**

## Statistische Signifikanz vs klinische Relevanz

Randomisierte Studie Kolonkarzinom:  
5FU versus Folfox:

- 5-Jahres Gesamtüberleben: 40% versus 60%:  
Differenz 20%: klinisch relevant!!

**Falls <146 Patienten/-innen in der Studie: keine signifikanter Unterschied**

## Statistische Signifikanz vs klinische Relevanz

Bei ungenuegend grosser Fallzahl sind auch klinisch relevante Unterschiede statistisch nicht signifikant.

## Relative Risikoreduktion

Die endoskopische Netzplastik (TEP) bei der Inguinalhernie verringert das Komplikationsrisiko um 50% im Vergleich zur offenen Methode.

## Relative Risikoreduktion

Die endoskopische Netzplastik (TEP) bei der Inguinalhernie verringert das Komplikationsrisiko um 50% im Vergleich zur offenen Methode.

**20% versus 10%: RELEVANT**

## Relative Risikoreduktion

Die endoskopische Netzplastik (TEP) bei der Inguinalhernie verringert das Komplikationsrisiko um 50% im Vergleich zur offenen Methode.

**2% versus 1%: NICHT RELEVANT**

## Relative Risikoreduktion

Beachten Sie immer auch die

**ABSOLUTE**

Differenz.

## Relative Risikoreduktion

Eine grosse relative Risikoreduktion kann mit einer kleinen absoluten Risikoreduktion einhergehen.

## Resultate

- Nicht verwertbar: **Papierkorb oder Letter to the Editor**
- Korrekt => **Schlussfolgerungen**

## Schlussfolgerungen

- Sind die Schlussfolgerungen auf die gefundenen Resultate abgestuetzt?
- Generalisierbarkeit ueber Ein- und Ausschlusskriterien hinaus?

## Zusammenfassung

- Journal
- Titel
- Abstract/Schlussfolgerungen
- Methodik
- Resultate
- Schlussfolgerungen

## Take Home Messages

- p-Wert: von Fallzahl abhängig
- Cave: statistische Signifikanz vs klinische Relevanz
- Confounding
- Absolute vs relative Risikoreduktion

## Take Home Messages

- Goldstandard: prosp. rand. Studie
- Auch prosp. rand. Studie hat Limitationen
- Prospektiv nicht zwingend besser als retrospektiv
- Retrospektiv nicht zwingend „low quality“

## Take Home Messages

- Ueben Sie!
- Statistik/Studien design

## Fazit

Wer in das Wesen der Statistik tief eindringt, wird zur Glückseligkeit und Ruhe der Weisheit kommen.

K. Peltzer

