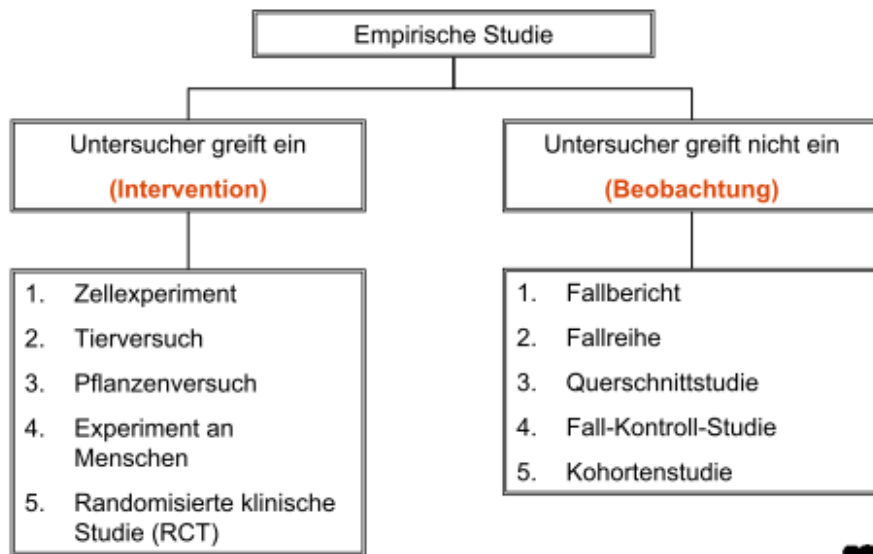


Epidemiologie

Studientypen.....	2
Häufigkeitsmaße	2
Häufigkeitsverteilung	5
Lagemaße	5
Streuungsmaße.....	6
Diagnostische Indizes	7
Probleme bei Diagnose-Studien	9
Klassifikationen.....	9
Methodik Studien	9
Phasen der Arzneimittelprüfung.....	10
Kontrollgruppen/Placebo	11
Randomisierung.....	11
Standardisierung	11
Cross-over-Design	12
Multizentrische kontrollierte klinische Studie.....	12
RCTs	12
Effektmaße	13
Konfidenzintervall.....	13
Statistische Tests	14
Fehler und p-Wert.....	14
Ereigniszeitanalyse(Überlebenszeit).....	14
Studienumfangplanung/ Auswertung	15
Studienprotokoll	16
Fallzahlplanung	16
Kohortenstudie	16
Fall-Kontroll-Studie	17
Odds-Ratio	17
Interpretation relatives Risiko und Odds Ratio.....	17
Vergleich Kohortenstudie und Fall-Kontroll-Studie	18
Confounding	18
Confounder	19
Bias	19
Evidenz-basierte Medizin.....	19
Klinische Informationssysteme	20
Formulierungen	20

Studientypen



prospektiv: Beobachtungseinheiten werden nach Studienbeginn fortlaufend (konsekutiv) rekrutiert, Daten im Beobachtungsverlauf erhoben.

retrospektiv: Informationen, Daten werden über zurückliegende Ereignisse erhoben.

Experimentelle Anlagen sind stets prospektiv, Beobachtungsstudien können, je nach Fragestellung, retrospektiv oder prospektiv sein.

empirische Studien:

- Ätiologische Studie (Risiko- & präventive Faktoren)
- Screeningstudie
- Diagnostische Studie
- Prognosestudie
- Therapiestudie
- Pathomechanistische Studie

fairer Vergleich „ceteris paribus“ (alles übrige gleich):

- zwei Gruppen, die hinsichtlich aller prognostischen Faktoren gleich sind, und sich nur hinsichtlich der Intervention/Exposition von Interesse unterscheiden
- zwei Gruppen dürfen sich allerdings in Faktoren, die nicht prognostisch wirksam sind, unterscheiden

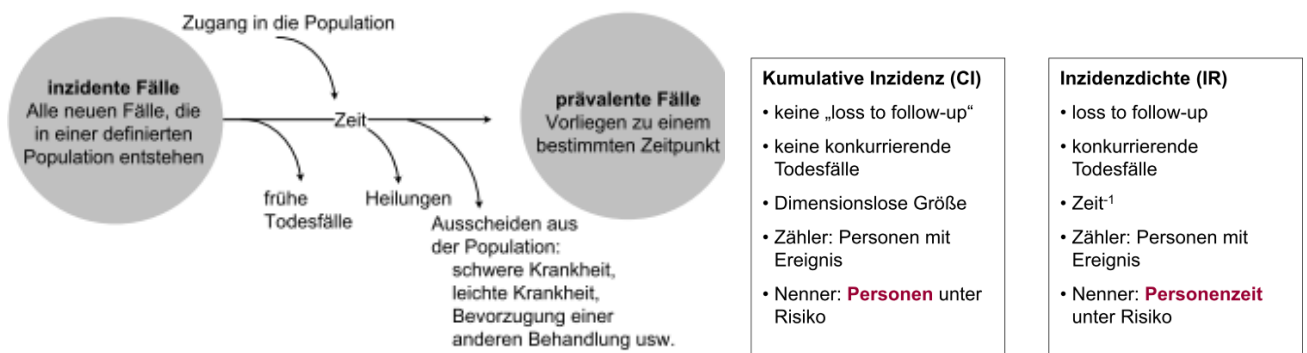
Häufigkeitsmaße

Kumulative Inzidenz: Anzahl der Neuerkrankten pro Jahr, bezogen auf die zu Beginn dieses (Erkrankungsrisiko) Zeitraums gesunde (nicht an der Zielkrankheit erkrankte) Bevölkerung

$$\text{Kumulative Inzidenz} = \frac{\text{Neuerkrankungsfälle}}{\text{alle Beobachteten u. Risiko}} = \frac{2}{5} = 40\%$$

Inzidenzrate / -dichte: bei instabiler Population

$$\text{Inzidenzrate} = \frac{\text{Neuerkrankungsfälle}}{\text{Personenjahre unter Risiko}} = \frac{2}{38 \text{ J}} = 0.05/\text{J} = 5 \text{ pro } 100 \text{ Personenjahre}$$



Inzidenzrate: loss-to-follow up relevant vorhanden, Interesse an Zeitlichkeit des Auftretens des Ereignisses
 z.B. Wegzug, Versterben an anderen Erkrankungen
 Kum. Inzidenz: loss-to-follow-up kaum vorhanden

- Inzidenzrate von 4/55 Personenjahre = 0,07273 Hcc-Fälle pro Jahr
 -> 7,273 pro 100 Personenjahre
 72,73 pro 1.000 Personenjahre
 727,3 pro 10.000 Personenjahre
 7273 pro 100.000 Personenjahre

Prävalenz: Anzahl der Kranken zu einem Zeitpunkt (Bestand an Kranken), bezogen auf die entsprechende Bevölkerung

Charakteristikum	Inzidenz	Prävalenz
Zähler	Während einer Zeitperiode neu auftretende Fälle unter Personen einer Gruppe, die anfangs frei von einer Krankheit waren	Alle Fälle, die bei einer einzigen Beobachtung oder Untersuchung einer Gruppe gezählt werden
Nenner	Alle prinzipiell empfänglichen Personen, die zu Beginn der Zeitperiode vorhanden sind	Alle untersuchten Personen, einschließlich Fälle und Nicht-Fälle
Zeit	Dauer der Zeitperiode	Einzelner Zeitpunkt
Wie gemessen	Kohortenstudie	Prävalenz-(Querschnitts-) Studie

Mortalität: Anzahl der Verstorbenen pro Jahr, bezogen auf die Bevölkerung zu Beginn des Zeitraums

Letalität: Anzahl der Verstorbenen, bezogen auf die an einer bestimmten Krankheit Erkrankten zu Beginn des Jahres
 = wenn erkrankt, wie groß die Wahrscheinlichkeit zu sterben?

Erkrankungsrate bei Autopsiefällen: Prävalenz

Mortalitätsmaße

- *rohe* Mortalitätsrate = ohne Berücksichtigung der Altersstruktur
- *altersspezifische* Mortalitätsrate = rohe Mortalitätsrate in einer Altersgruppe
- *altersstandardisierte* Mortalitätsrate = künstlich errechnete Mortalitätsrate nach Standardisierung (Anpassung an eine externe Standard - Altersstruktur)

- *altersadjustierte* Mortalitätsrate = Adjustierung hinsichtl. des Alterseinflusses mit einem statistischen Regressionsmodell

Datenqualität

- die **Validität** (sie ist ein Maß dafür, ob die Daten das messen, was sie messen sollen) und
- die **Reliabilität** (Kriterium dafür, ob sich bei wiederholter Messung unter gleichen Bedingungen auch gleiche Werte ergeben).

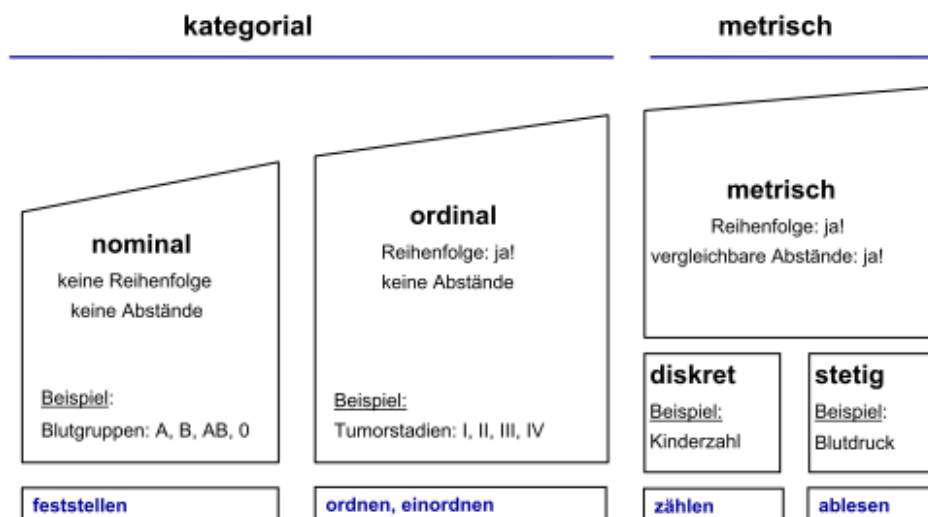
Zwei Typen von Merkmalen:

metrische Merkmale

- metrische gemessene (stetige, kontinuierliche) oder
- ausgezählte (diskrete) Merkmale

kategoriale Merkmale

- nominal skalierte Merkmale
 - mit zwei Werten (binäre oder dichotome Merkmale)
 - mit mehr als zwei Werten (polychotome Merkmale)
- ordinal skalierte Merkmale



- Beispiele:
- **nominal:** Blutgruppen, Geschlecht
 - **ordinal:** TNM, Lebensqualität, NYHA
 - **metrisch (diskret):** Kinderzahl
 - **metrisch (stetig):** Alter, Blutdruck, Kreatinkinase, Leukozyt

Outcome

klinischer Endpunkt (**Outcome**): ein für den Patienten erfahrbarer klinischer Zustand (Endzustand) mit Relevanz für den Krankheitsverlauf

z.B. Tod, Krankheit, Beschwerden, Behinderung, Lebensqualität

Surrogat-Endpunkte:

- der eigentliche Endpunkt schwer zu messen
- eine Messung ethisch nicht vertretbar
- zeitlich zu lange Beobachtungszeiten

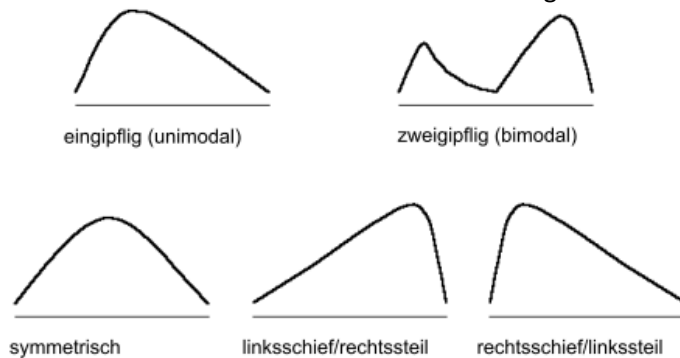
-> als Ersatz werden Zielgrößen als Endpunkte benutzt, die mit dem eigentlichen Endpunkt statistisch und inhaltlich assoziiert sind

Beispiele: Zucker im Urin anstelle von Zucker im Blut;
Cholesterinwert anstelle des Auftretens eines kardiovaskulären Krankheitsereignisses

Häufigkeitsverteilung

= zahlenmäßige oder graphische Darstellung absoluter oder relativer Häufigkeiten

graphische Darstellung: Histogramm (Säulendiagramm)
seltener Polygonzug (Linienzug aus Geradenstücken)
kumulative Häufigkeitsverteilung



Lagemaße

arithmetischer **Mittelwert**: $\bar{x} = \frac{\sum x_i}{n}$

Median (Zentralwert) **z**: ist definiert als Wert in der Mitte in der geordneten Reihe der Messwerte. Das heißt, dass 50 % der Werte größer sind als der Median einer Reihe und 50 % kleiner.
 $z = n+1 / 2$

Modalwert (der häufigste Wert):

oft nicht eindeutig

nur bei gruppierten Daten sinnvoll

geeignet für alle Skalenniveaus

Median ist dem **Mittelwert** vorzuziehen bei

- asymmetrischer Verteilung
- Verteilung mit Ausreißern
- erfordert ordinale Daten

Mittelwert ist „genauer“ als der **Median**

- wenn symmetrische Verteilung der Daten vorliegt (dann sind aber beide Größen ähnlich!)
- erfordert metrische Daten

Perzentilwerte: Rangreihe der Werte

Die **Perzentile** $P_{h\%}$ bezeichnen diejenigen Skalenwerte, bis zu denen $h\%$ der Beobachtungswerte liegen. Besonders häufig genutzt werden die Perzentile P_5 , P_{10} , P_{25} , P_{75} , P_{90} , P_{95} .

Interpretiert werden Perzentile so, dass zum Beispiel P_5 den Wert darstellt, für den die kumulative Häufigkeit, dass Beobachtungswerte kleiner oder gleich diesem Wert vorkommen, 5 % beträgt.

Quartile = spezielle Perzentile mit: P_{25} , P_{50} , P_{75}

Streuungsmaße

Wie stark streuen die Messwerte um ihr „Zentrum“ (=Lagemaß) herum?

Quellen der Variabilität

- biologische Variabilität (intraindividuell, interindividuell)
- messtechnische Variabilität
- Untersuchervariabilität
- Zufall

Standardabweichung s, Varianz s²

$$s^2 = \frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x})^2$$

$$s = \sqrt{s^2}$$

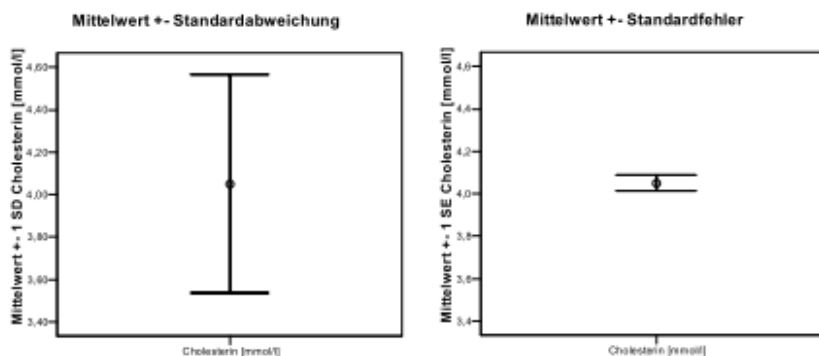
Standardfehler s : zur Charakterisierung der Variabilität des Mittelwertes

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

Graphische Darstellungen mit Hilfe der Lage- und Streuungsmaße

Bei der graphischen Darstellung von Daten mit Fehlerbalken gibt es zwei unterschiedliche Möglichkeiten:

Mittelwert ± Standardabweichung	Mittelwert ± Standardfehler
Die Darstellung unterstützt den Gruppenvergleich durch Beschreibung der Variabilität der Einzelwerte.	Die Darstellung unterstützt den Gruppenvergleich durch die Beschreibung der Variabilität des Mittelwertes.



Streuungsmaß (Range) = $x_{\max} - x_{\min}$

Normalverteilung

- kontinuierliche Verteilung
- hängt ab von: Erwartungswert, Standardabweichung

Grundgesamtheit

- Erwartungswert μ
- Varianz σ^2

→

Stichprobe

- Mittelwert \bar{x} quer
- empirische Varianz s^2

- Interpretation: Flächen unter der Normalverteilungskurve können als Wahrscheinlichkeiten interpretiert werden.

- σ -Bereich = 68% 2 σ -Bereich = 95% 3 σ -Bereich = 99,7%

Korrelation

Korrelationskoeffizient r (Pearson): Maß der Stärke des Zusammenhangs.

Der Koeffizient ist eine dimensionslose Größe. Er liegt zwischen $r=-1$ und $r=1$.

Dabei bedeutet

- für ein negatives Vorzeichen ($r<0$) von r eine Beziehung „je größer x , desto kleiner wird y “ (gegenläufige Beziehung)
- für ein positives Vorzeichen ($r>0$) von r eine Beziehung „je größer x , desto größer wird y “ (gleichläufige Beziehung).

Regressionsgerade: $y = \beta_0 + \beta_1 \cdot X$

Nachteile der Statistischen Definition von Normal (Perzentilen)

1. Exakt gleiche Häufigkeit von so definierten Erkrankungen (upper limit syndrome of nondiseased)
2. Populations-abhängige Definition
3. Mehrere unabhängige Tests an Gesunden senkt Wahrscheinlichkeit als Gesunder eingestuft zu werden

“Gold“Standard der diagnostischen Abklärung

Perfektes diagnostisches Verfahren

Meßfehlerfrei! (Sensitivität = Spezifität = 100%)

Ausnahme in der Medizin, besser: **Referenzstandard**

Nachteile beim Referenzstandard-Einsatz

Invasivität & Risiken

Kosten, Zeitaufwand

Begrenzte Verfügbarkeit

Ersatzweiser Einsatz eines diagnostischen Verfahrens mit Fehlern

Extra-Unsicherheit

Diagnostische Indizes

Diagnostischer Test	"Wahrheit": Krankheit liegt		Sn = $d/(b+d)$	Sp = $a/(a+c)$
	nicht vor	liegt vor		
unauffällig	RN	FN	PPV = $d/(c+d)$	NPV = $a/(a+b)$
auffällig	FP	RP		

R: richtig, F: falsch, N: negativ, P: positiv

Positiver prädiktiver Wert (PPV)

...gegeben, dass der Test auffällig ausfällt, wie hoch ist dann die Wahrscheinlichkeit, dass die Erkrankung tatsächlich vorliegt?

Negativer prädiktiver Wert (NPV)

...gegeben, dass der Test unauffällig ausfällt, wie hoch ist dann die Wahrscheinlichkeit, dass die Erkrankung tatsächlich nicht vorliegt?

Sensitivität (Sn)

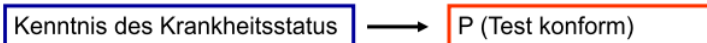
...gegeben, dass die Erkrankung tatsächlich vorliegt, wie hoch ist dann die Wahrscheinlichkeit, dass der Test auffällig ausfällt?

Spezifität (Sp)

...gegeben, dass die Erkrankung tatsächlich nicht vorliegt, wie hoch ist dann die Wahrscheinlichkeit, dass der Test unauffällig ausfällt?

Sensitivität und Spezifität

- Testgüteeigenschaften
- Unabhängig von der Vortestwahrscheinlichkeit des Vorliegens einer Erkrankung



Prädiktive Werte

- Vorhersagewahrscheinlichkeiten
- Abhängig von: Sn, Sp, Vortestwahrscheinlichkeit des Vorliegens einer Erkrankung



- Bei gegebener Sn und Sp ist der PPV von der Prävalenz abhängig.

$$PPV = \frac{P(D^+) \cdot Sn}{P(D^+) \cdot Sn + P(D^-) \cdot (1 - Sp)}$$

$$NPV = \frac{P(D^-) \cdot Sp}{P(D^-) \cdot Sp + P(D^+) \cdot (1 - Sn)}$$

- über Anamnese wird die Vortestwahrscheinlichkeit gesteuert

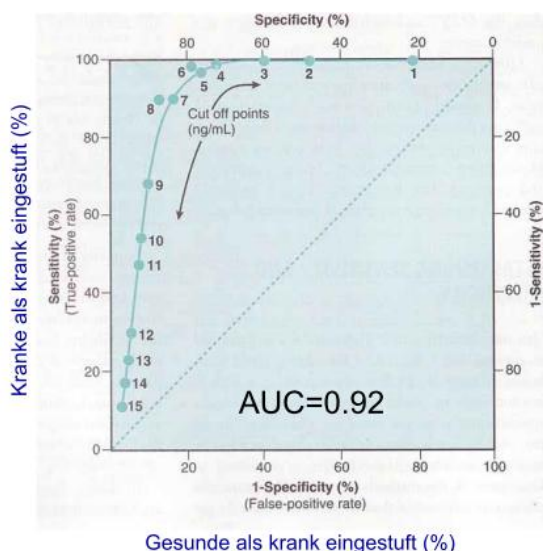
hohe Sensitivität nötig: bei schweren Nachteilen durch Übersehen der Krankheit (lebensbedrohlich, ansteckend)

hohe Spezifität nötig: wenn falsch-pos. Test dem Patienten physisch, emotional oder finanziell schaden kann, Vorsorgeuntersuchung

PPV: je spezifischer, desto besser PPV; hohe Prävalenz -> PPV hoch

NPV: je sensitiver, desto besser NPV

ROC-Kurve



- je größer die Fläche unter der Kurve (AUC), desto besser die Güteeigenschaften des Tests
- optimaler Punkt zur Minimierung der Fehler: Punkt der am weitesten oben links liegt

- Falsch-Positive = 1-Spezifität
- Falsch-Negative = 1-Sensitivität

- Cut- Point ↓ = Sensitivität ↑, Spezifität ↓

Probleme bei Diagnose-Studien

Fehlender Referenzstandard

- Experten Panel
- Chronische Erkrankungen: Nachverfolgung der Studienteilnehmer

Review Bias

- Test wird in Kenntnis des Referenzstandards befundet (vice versa)

Spektrum Bias

- Erkrankte nicht repräsentativ (z.B. asympt/symp. Frauen & Mammographie)

Inkorporations Bias

- Testergebnis ist Teil der diagnost. Kriterien für Referenzstandard

- Erhöhung Sn => Verminderung Sp (& vice versa)
- Erhöhung PPV => Verminderung NPV (& vice versa)

Wahl des Cutoffs

- Gefährlichkeit der Erkrankung
- Behandlungsmöglichkeit der Erkrankung
- Nebenwirkungen der Behandlung
- Psychologische Auswirkungen eines falsch-positiven Alarms (z.B. Krebs, HIV)
- Gesellschaftliche Folgen eines falsch-negativen Tests (z.B. Ebola etc.)
- Rahmenbedingung (Screening, Klinik)
- Kosten
- Patientenpräferenz

Klassifikationen

- Bedeutung:
1. Kommunikationshilfe
 2. Prognose
 3. Entscheidungshilfe
 4. Unterstützung des Qualitätsmanagements
 5. Gesundheitsberichtserstattung
 6. Statistische Auswertung / Studien
 7. Abrechnung

Methodik Studien

numerische Methode: therapeutisches Mittel kann nur dann mit sicherem Urteil und Aussicht auf Erfolg eingesetzt werden, wenn **Statistik** Wirkung zeigt

- Randomisierte klinische Studie (RCT)**
- prospektiv
 - streng zufällige Zuteilung
 - kontrollierte Durchführung
 - Studienprotokoll

Phasen der Arzneimittelprüfung

Phase	Inhalt	Teilnehmer
I	Pharmakokinetik, Bioverfügbarkeit, Verträglichkeit	Gesunde
II	Erprobung der Wirksamkeit, Dosisfindung	Wenige Kranke
III	Strenge Erprobung der Wirksamkeit (RCT)	Viele Kranke
Arzneimittelzulassung		
IV	Anwendungsbeobachtungen, Postmarketing-Studie	Bevölkerung

Seltene Nebenwirkungen
Spät auftretende Nebenwirkungen

Phase I Studien

Erste Studien einer neuen aktiven Substanz beim Menschen, oftmals bei gesunden Probanden. Das Ziel ist, eine vorläufige Bewertung der Unbedenklichkeit der Anwendung und eine erstmalige Beschreibung des pharmakokinetischen/pharmakodynamischen Profils der aktiven Substanz beim Menschen zu erhalten.

Phase II Studien

Das Ziel ist, eine Aktivität nachzuweisen und eine Bewertung der Unbedenklichkeit der aktiven Substanz nach Kurzzeitanwendung vornehmen zu können. Die Studien werden bei einer begrenzten Anzahl von Patienten und oft, in einem späteren Entwicklungsstadium, als vergleichende Studien durchgeführt.

In dieser Entwicklungsphase sind Dosisfindungsstudien durchzuführen und eine Abklärung der Beziehung zwischen Dosis und Wirksamkeit, um eine optimale Ausgangssituation für die Planung größerer therapeutischer Studien zu erhalten.

Phase III Studien

Studien an größeren Patientengruppen mit dem Ziel, die Risiko-Nutzen-Abwägung nach kurz- und längerfristiger Gabe des Wirkstoffes zu ermöglichen und den therapeutischen Wert im Vergleich zu therapeutischen Alternativen zu bestimmen.

Das qualitative und quantitative Muster für das Auftreten häufiger unerwünschter Ereignisse muss abgeklärt werden und Besonderheiten des Arzneimittels müssen aufgefunden werden (z. B. Interaktionen).

Die Studienanordnung sollte vorzugsweise ein doppelblindes, randomisiertes Design aufweisen. Generell gilt, dass die Bedingungen der Realität so nahe wie möglich kommen sollten.

Phase IV Studien

Untersuchungen nach Zulassung des Fertigarzneimittels; allerdings besteht über die Definition dieser Phase keine vollständige Einigung. Prüfungen in Phase IV werden durchgeführt auf der Grundlage der im Zulassungsbescheid festgelegten Bedingungen z. B. Überwachung

nach dem In-Verkehr-Bringen, Bewertung des therapeutischen Werts oder von Strategien. Entsprechend den Umständen machen Phase-IV Studien Studienbedingungen (einschließlich eines Prüfplanes) erforderlich, wie sie für Studien vor Einführung auf dem Markt beschrieben sind.

Kontrollgruppen/Placebo

Arten von Kontrollgruppen:

- keine Behandlung
- usual care
- Placebo
- aktive andere Substanz (A vs. B)
- andere Dosis als Testsubstanz (A1vs A2)
- (historische Vergleiche)

Placebo: Scheinmedikament

Standardpräparat: Medikament, dessen Wirksamkeit unbestritten ist (in der Regel in einer kontrollierten klinischen Studie geprüft)

Placebo darf nur eingesetzt werden

- falls es noch keine anerkannte Therapie gibt
 - bei Krankheiten, die nicht sofort behandelt werden müssen
- Placebo hat auch immer Effekt (Suggestivwirkung) -> Effekt des neuen Medis wird nicht überschätzt

Randomisierung

Strukturgleichheit: Randomisierung

Behandlungs-/Beobachtungsgleichheit: Standardisierung / Maskierung

Gleichermaßen für alle Einflussfaktoren, gleich ob beobachtet
prinzipiell beobachtbar, aber faktisch nicht beobachtet
unbeobachtbar

Ziele:

- Vermeidung einer Verzerrung (Sicherung der Vergleichbarkeit der Gruppen)
- Balanciertheit (ausgeglichene Anzahl von Patienten in den Gruppen)

Methoden:

- totale Randomisierung ohne Ausgleich
- totale Randomisierung mit Ausgleich (permutierte Blocks)
- stratifizierte Randomisierung (innerhalb der Schichten prognostischer Faktoren)

Ethische Zulässigkeit einer zufälligen Zuteilung

- es existiert bislang kein Nachweis, dass eine der Behandlungen besser ist
 - alle Studienmaßnahmen sind ethisch vertret- und zumutbar
 - Patientenaufklärung und Einverständnis („informed consent“) liegt vor
- Zuordnung darf nicht vorhersehbar sein! (allocation concealment)

Standardisierung

Standardisierung der Begleitbehandlungen und aller Mess-, Befund- und Beobachtungsvorgänge

Ziel: Vermeidung von Bias nach Randomisierung

Bias: systematische Verzerrung zwischen den Behandlungsarmen

Idee: Vermeidung von **Informations-Bias**

Verblindung

- einfach-blind: Patienten werden hinsichtlich Medikation verblindet (würde sich sonst vll anders verhalten)
- doppel-blind: Patienten und nachbeobachtende Ärzte werden hinsichtlich Medikation verblindet (Arzt untersucht anders, kann nicht neutral sein)
- dreifach-blind: Patienten, nachbeobachtende Ärzte und auswertende Biometriker werden hinsichtlich Medikation verblindet

Cross-over-Design

Patienten erhalten sowohl die Studien- als auch die Kontrolltherapie, aber in verschiedener Reihenfolge.

nur anwendbar bei chronischen Krankheitsverläufen und Therapien, von denen man kurzfristige Linderungen, aber keine Heilung erwartet (z.B. chronischer Schmerz)

Vorteil: Patient ist eigene Kontrolle (geringere Variabilität, kleinerer Studienumfang); höhere Mitwirkungsbereitschaft

Nachteil: Nachwirkungseffekt der zuerst gegebenen auf die nachfolgende Therapie (carry-over-Effekt); zur Vermeidung wird wash-out-Phase dazwischengeschaltet

Multizentrische kontrollierte klinische Studie

- f in mehreren Kliniken oder Arztpraxen
- f bei relativ seltenen Krankheiten oder seltener Indikationsstellung ausreichend große Fallzahlen
- f jedes Zentrum muss eine Mindestzahl von Fällen beisteuern
- f Schaffung einer unabhängigen Studienzentrale (z. B. biometrisches Institut), die mit dem Monitoring der Studie beauftragt wird

Studiendesign – Was muss ich beurteilen?

- f Welche Phase einer klinischen Prüfung?
- f Mono- oder multizentrisch?
- f Wie viele Therapiearme, welche Kontrollgruppe(n)?
- f Parallelgruppen- oder Crossover-Design?
- f Behandlungszuweisung nach Arzt-/ Patientenentscheidung oder randomisiert ?
- f Behandlung / Bewertung: Offen oder verblindet(einfach oder doppelt)?

RCTs

Das Ergebnis wird gemessen an im Vorfeld definierten Zielgrößen (Endpunkten).

- Hauptzielkriterium (primärer Endpunkt)
- Nebenzielkriterien (sekundäre Endpunkte)

Problem: Multiple Endpunkte:

Multiple primäre Endpunkte erfordern komplexe biometrische Verfahren zur Ermittlung des Studienumfangs und zur Auswertung (multivariate statistische Analysen) – oft statistisch noch ungelöste Probleme.

Effektmaße

abhängige	vs.	unabhängige Gruppen
abhängige Gruppen (paired samples)		unabhängige Gruppen (independent samples)
Beispiele		
vorher - nachher		Männer vs. Frauen
links - rechts		Kranke vs. Gesunde
Zeitpunkt 1 - Zeitpunkt 2		Verum vs. Placebo
Zwilling 1 - Zwilling 2		(versch. Patientengruppen)
Verum vs. Placebo (cross over)		

Kontinuierliche Zielgröße (Messwert)

Mittelwertdifferenz $\bar{X}_1 - \bar{X}_2$

Mediandifferenz $z_1 - z_2$

dichotome Zielgröße (Ereignis eingetreten, nicht eingetreten)

Risikodifferenz $RD = f_1 - f_2$ (Prozentpunkte)

Relatives Risiko $RR = \frac{f_1}{f_2}$

Odds Ratio (relative Risikoquote;
Chancenverhältnis) $OR = \frac{f_1 / (1 - f_1)}{f_2 / (1 - f_2)}$

	D+	D-	
E+	a	b	n_1
E-	c	d	n_2
	$f_1 = \frac{a}{n_1}$	$f_2 = \frac{c}{n_2}$	

„Die **Risikodifferenz** der Gesamtmortalität zwischen Simvastatin- und Placebogruppe liegt mit 95%iger Wahrscheinlichkeit zwischen -5,0 und -1,6 PP.“

Relative Risikoreduktion: $RRR = 1 - RR$

„Das Sterberisiko nach Myokardinfarkt/Angina pectoris wird durch Statin-Behandlung um (relativ) 29 % gesenkt (Konfidenzintervall 15 %ige bis 41 %ige Reduktion)“

Number Needed to Treat (Anzahl zu beobachtender Patienten, um einen Patienten zu retten; bei einem Patienten des Zielereignis zu vermeiden) $NNT = \frac{1}{-RD}$

„Es müssen 30,3 Patienten nach Myokardinfarkt/Angina pectoris mit Statinen (im Vergleich zu Placebo) behandelt werden, um einen Todesfall zu verhindern.“

Konfidenzintervall

Angabe eines Schwankungsbereichs, eines Bereichs, in dem der Populationswert des Effektes mit einer bestimmten Wahrscheinlichkeit erwartet werden kann.

$CI_{1-\alpha} =$ (untere Intervallgrenze ; obere Intervallgrenze)

„Das Konfidenzintervall schließt den unbekannten Effekt mit einer Vertrauenswahrscheinlichkeit $1 - \alpha$ ein.“

Schärfe und Sicherheit von Konfidenzaussagen:

Je schmaler das Intervall, desto schärfer die Konfidenzaussage.

Je höher die Sicherheitsforderung für die Konfidenzaussage, desto breiter das Intervall.

- je größer die Stichprobe, desto kleiner die Länge des KI
- je größer die gewünschte Sicherheit, desto größer das KI

- je größer die Streuung der Werte, desto größer das KI

Statistische Tests

Wissenschaftliche Hypothese:

Das neue Medikament führt zu einer deutlicheren Veränderung des Blutdruckes als Placebo.

Statistische Hypothese:

Nullhypothese:

Die mittlere Veränderung des Blutdruckes bei Gabe des neuen Medikaments unterscheidet sich nicht von der mittleren Veränderung des Blutdruckes bei Placebo-Behandlung.

Alternativhypothese:

Die mittlere Veränderung des Blutdruckes bei Gabe des neuen Medikaments unterscheidet sich von der mittleren Veränderung des Blutdruckes bei Placebo-Behandlung.

Fehler und p-Wert

Irrtumswahrscheinlichkeit α :

- f Signifikanzniveau
- f Fehler 1. Art: Fehlentscheidung, die Nullhypothese abzulehnen, obwohl sie zutrifft.

Irrtumswahrscheinlichkeit β :

- f Fehler 2. Art: die Nullhypothese beizubehalten, obwohl H A gilt.

Die Größe $1 - \beta$ wird Teststärke (**power**) genannt. Sie beschreibt die Macht eines Tests, einen vorhandenen Unterschied auch zu erkennen.

Durch das Festlegen des α -Fehler vor Beginn des Experiments haben wir diesen eingeschränkt, d.h. wir machen diesen nur mit 5% Wahrscheinlichkeit.

Wenn der Test also die Nullhypothese ablehnt, machen wir mit (mindestens) 95%-iger Wahrscheinlichkeit keinen Fehler.

	Nullhypothese wahr	Alternativhypothese wahr
Nullhypothese akzeptiert	Richtige Entscheidung	Typ II Fehler (Falsch Negative)
Nullhypothese abgelehnt	Typ I Fehler (Falsch Positive)	Richtige Entscheidung

p-Wert:

- gibt die Wahrscheinlichkeit an, vorliegende oder extremere Versuchsvorgänge zu beobachten wenn die Nullhypothese zutrifft
- quantifiziert die Wahrscheinlichkeit, dass das ermittelte Testergebnis aufgrund des Zufalls zustande kommt
- $p \leq \alpha$ = signifikant -> Nullhypothese verwerfen
- sagt nichts über Ursachen, über wissenschaftliche oder praktische Relevanz
- p-Wert sollte immer zusammen mit jeweiligen Konfidenzintervallen angegeben werden
- signifikant = durch Zufall nur mit bestimmter geringer Wahrscheinlichkeit zustande gekommen

Ereigniszeitanalyse(Überlebenszeit)

analysiert die Zeit bis zum Eintreten eines Ereignisses

- Tod, Krankheit, Rückfall, Heilung
- Abstoßung eines Transplantats
- Überschreiten eines Grenzwertes

Gründe für das Abbrechen der Beobachtung:

- Ende der individuellen Beobachtungszeit
- Studienende
- Wegzug, non-compliance, andere Todesursache (Unfall)

Beobachtungszeiten (Überlebenszeiten), deren Beobachtung abgebrochen wird, nennen wir zensiert (genauer: rechts zensiert).

rezidivfreies Intervall: Zeit bis zum Eintreten eines Rezidivs

Remissionszeit: Zeit der Zurückdrängung von Krankheitserscheinungen (bis zum Wiederauftreten der Erscheinungen)

Expositionszeit: Zeit von Beginn einer Exposition bis zum Krankheitseintritt

Beobachtungszeit: Zeit bis zum Ende der Beobachtung (in Monaten, Jahren)

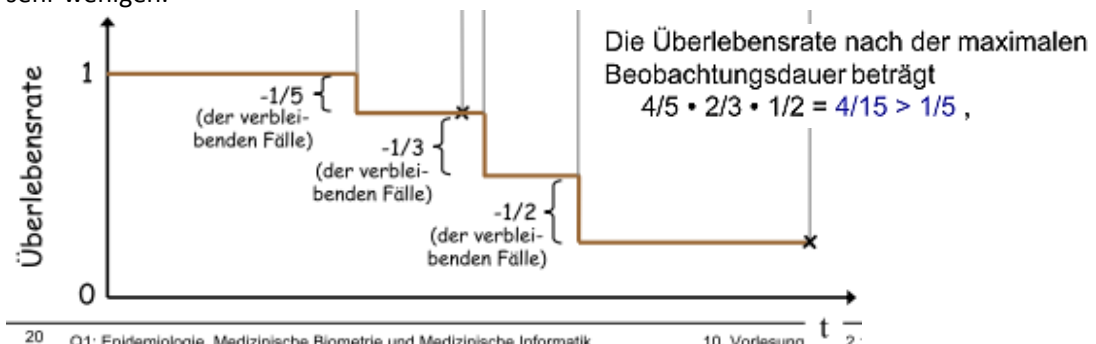
Überlebensstatus: Status am Ende der Beobachtung

Zielereignis eingetreten 1

Zielereignis nicht eingetreten (zensiert) 0

Kaplan-Meier-Schätzung von Überlebenskurven

- Schätzung der Überlebenskurve (Absterbekurve) aus den individuellen Überlebenszeiten (Sterbezeitpunkten)
- Grafische Darstellung in Form einer Treppenfunktion mit Stufen zu jedem Sterbezeitpunkt.
- Geeignet auch für kleine Studienumfänge.
- Überlebenskurven sind am aussagekräftigsten im linken Bereich (zu Beginn): Zu Beginn beruhen die Schätzungen noch auf beinahe allen Patienten, am Ende (rechts) nur noch auf sehr wenigen.



- **Log-Rank-Test** zum Vergleich von Kaplan-Meier-Kurven (Chi-Quadrat-Test -> p-Wert)
- Einfluss von mehreren Faktoren auf die Überlebensfunktion: **Cox-Regression (Hazardratio)**
- **Hazard-Ratio:** wie relatives Risiko, aber verschiedene Follow-Up-Zeiten werden berücksichtigt

Studienumfangplanung/ Auswertung

- Faustregel 1: Je geringer der nachzuweisende klinisch relevante Unterschied gefordert wird, desto höher die Fallzahl der Studie!
- Faustregel 2: Je größer die Streuung in einem stetigen primären Endpunkt, desto höher die Fallzahl der Studie!
- Faustregel 3: Je geringer α - und β -Fehler gefordert werden, desto höher die Fallzahl der Studie!
- Also: je sicherer man sein will, desto mehr Fälle braucht man!

Auswertungspopulationen

Intention-to-treat	Per-protocol
Vorteile: <ul style="list-style-type: none"> Strukturgleichheit bleibt erhalten bildet Behandlungsrealität besser ab Nachteil: <ul style="list-style-type: none"> tatsächlicher Therapieeffekt wird oft unterschätzt 	Vorteile: <ul style="list-style-type: none"> klare Zuordnung der Behandlung bildet Wirkungsmechanismen ab Nachteile: <ul style="list-style-type: none"> Strukturgleichheit geht ggf. verloren multivariate Statistik Fallzahl wird geringer

Therapieabbrecher (drop-outs, losses):

Patienten gehen im Verlauf der Prüfung irgendwie verloren, keine Information über das Zielkriterium

Studienausschlüsse (withdrawals):

Patienten werden während der klinischen Prüfung oder aus der Auswertung ausgeschlossen.

Studienprotokoll

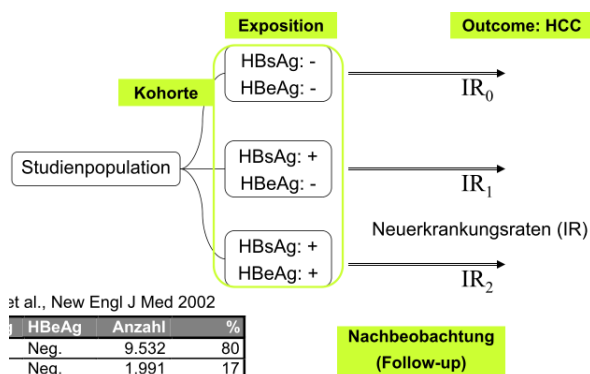
- sollte enthalten:
- Methoden
 - Ein- und Ausschlusskriterien
 - Outcome definieren, auch sekundäre
 - Art der Randomisierung
 - Dauer
 - Confounder
 - Nebenwirkungen
 - wie standardisiert man die Sache angeht
 - Auswertungsart
 - Subgruppenanalyse

= gesamter Plan, jeder Schritt!

Fallzahlplanung

- Informationen, die benötigt werden: Prävalenzen, Fehler 1. (α) und 2. Art (β)
- seltene Erkrankung -> mehr einschließen: Prävalenz \downarrow -> Fallzahl \uparrow
- größeren β -Fehler zulassen (Power geringer) -> kleinere Gruppe reicht: Fehler \uparrow -> Fallzahl \downarrow

Kohortenstudie



Fall-Kontroll-Studie

Aus der Population unter Risiko werden Patienten, die die Krankheit haben, und eine Gruppe sonst ähnlicher Menschen, die nicht an der interessierenden Krankheit leiden, ausgewählt.

Die Untersucher blicken dann zeitlich zurück, um die Häufigkeit von bestimmten Risikofaktoren in beiden Gruppen zu erfassen.

Diese Daten werden dann genutzt, um das Odds Ratio der Krankheit, das mit der interessierenden Exposition in Zusammenhang steht, zu schätzen.

- Ereignisse = Kohortenstudie | Fälle = Fall-Kontroll-Studie
- Risiko-Set-**Sampling**-Prozedur: Gewinnen von Kontrollen, parallel zum Auftreten des Falles (gleiche Zeit), denn jedes Mal ist eine andere Zusammensetzung von Personen unter Risiko!
- seltene Erkrankung: Fall-Kontroll-Studie (Kohortenstudie wäre insuffizient, sehr große Kohorte nötig)
- häufige Erkrankung: Kohortenstudie
- Personenzeit fehlt in Fall-Kontroll-Studie

Odds-Ratio

Exponiert	Kontrollen (N)	Fälle (N)
Nein	B ₀	A ₀
Ja	B ₁	A ₁

$$OR = \frac{\text{Expositions – Odds(Fälle)}}{\text{Expositions – Odds(Kont)}} = \frac{\frac{A_1}{B_1}}{\frac{A_0}{B_0}}$$

Nulllage = 1,0 | >1,0 : wenn exponiert = erhöhtes Risiko (Odds, Chance) | <1,0 Exp.=präventiver Effekt

Interpretation relatives Risiko und Odds Ratio

- Bei einem RR/OR von 1 besteht kein Zusammenhang zwischen der Exposition und der Erkrankung.
- Bei einem RR/OR von kleiner 1 wirkt die Exposition als **Schutzfaktor**, also protektiv.
- Bei einem RR/OR von größer 1 wirkt die Exposition als **Risikofaktor**.

Interpretation des relativen Risikos / Odds Ratio

Das Odds Ratio lässt sich tendenziell wie ein relatives Risiko (RR) interpretieren (unter bestimmten Bedingungen stimmen beide Effektmaße tatsächlich überein).

- **RR/OR =1:** Das Erkrankungsrisiko/Odds Ratio für Exponierte ist genauso hoch wie für Nicht-Exponierte, d.h. zwischen Exposition und Erkrankung liegt kein Zusammenhang vor.
- **RR/OR =0,33:** Das Erkrankungsrisiko ist für Exponierte um 2/3 kleiner als für Nicht-Exponierte.
Oder
Das Erkrankungsrisiko ist für Exponierte im Vergleich zu den Nichtexponierten um 67% niedriger.

RR bzw. OR=2: Es besteht ein Zusammenhang zwischen Exposition und Erkrankung. Das Erkrankungsrisiko/Odds Ratio für Exponierte ist doppelt so hoch wie für Nicht-Exponierte.

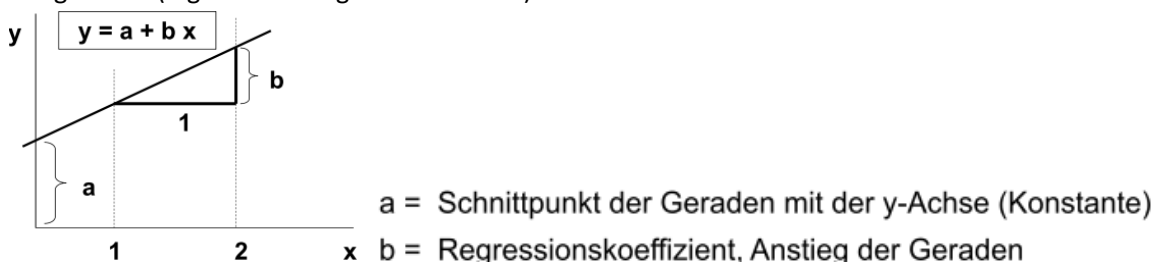
Vergleich Kohortenstudie und Fall-Kontroll-Studie

Kohortenstudie	Fall-Kontroll-Studie
<ul style="list-style-type: none"> - bei häufigen Erkrankungen = genug Ereignisse <p>Vorteile:</p> <ul style="list-style-type: none"> - hohe Datenqualität - Risiken können direkt berechnet werden (Personenjahre) - mehrere Erkrankungen gleichzeitig untersuchbar <p>Nachteile:</p> <ul style="list-style-type: none"> - viele Teilnehmer nötig - langwierig - teuer - riesige Kohorte bei seltenen Erkrankungen benötigt - mögliche Verzerrung bei langer Dauer 	<ul style="list-style-type: none"> - bei seltenen Erkrankungen <p>Vorteile:</p> <ul style="list-style-type: none"> - billiger - schneller <p>Nachteile:</p> <ul style="list-style-type: none"> ▪ Temporalität (Ursache muss vor Wirkung aufgetreten sein) ▪ Selektionsbias (niedrige Teilnahmerate und z.B. mehr Exponierte = Verschiebung) ▪ Informationsbias (Fehlklassifikationsbias) durch soziale Erwünschtheit, Erinnerungsprobleme, individuelles Kausalitätsproblem (man denkt mehr drüber nach)

Confounding

Confounding: Verzerrung des interessierenden Effekts durch Kofaktoren

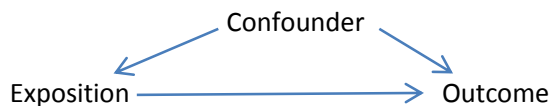
- Prävention: - **Randomisation**
- (Vermeiden) - **Restriktion**, z.B. nur Männer (Geschlecht)
- **Matching** (z.B. nur Leute mit gleichem BMI)
- hinterher rausrechnen: - **Stratifizierung** (Schichtung für den Confounder, z.B. Alter)
- (Korrektur)
- **Regression** (Adjustierung)
- **Standardisierung** (Neugewichtung der Strata -> re- gewichtetes Risiko (gleiche Besetzung Test- und Kontrollgruppe))
- **Null-Effekt** = Exposition von Interesse hat gar keinen Einfluss auf Outcome
- kein Confounding: 2 Gruppen unterscheiden sich nur hinsichtlich der Exposition von Interesse (ceteris paribus), alle anderen Einflussfaktoren gleich
- Regression (log-lineares Regressionsmodell)



Confounder

Nach der klassischen Definition des Confoundings (Effektvermischung) müssen folgende Bedingungen erfüllt sein, damit ein Faktor als Confounder wirken kann, das bedeutet, unerwünschte Effektvermischung produziert:

1. Der Faktor muss einen Einfluss auf die Exposition von Interesse haben.
2. Der Faktor muss einen Einfluss auf das Outcome von Interesse haben (auch bei denjenigen die nicht gegenüber der Exposition von Interesse exponiert sind).
3. Der Faktor darf nicht Mediator zwischen der Exposition von Interesse und dem Outcome von Interesse sein.



Bias

Bias: Fehler während der Planung/Durchführung einer Studie

- **Selektionsbias**: niedrige Teilnahmerate und z.B. mehr Exponierte = Verschiebung
- **Informationsbias**: (Fehlklassifikationsbias) durch soziale Erwünschtheit, Erinnerungsprobleme, individuelles Kausalitätsproblem (man denkt mehr drüber nach)

Bias bei Tests (z.B. Sensitivität weicht von Wahrheit ab)

- | | |
|----------------------------|--|
| <i>Review Bias</i> | Test wird in Kenntnis des Referenzstandards befundet (vice versa) -> weiß wonach er gucken muss |
| <i>Spektrum Bias</i> | Erkrankte nicht repräsentativ (z.B. asympt/symp. Frauen & Mammographie) -> nur total Kranke und total Gesunde -> dem Test wird es leicht gemacht |
| <i>Inkorporations Bias</i> | Testergebnis ist Teil der diagnost. Kriterien für Referenzstandard |

Evidenz-basierte Medizin

EBM ist der gewissenhafte, ausdrückliche und vernünftige Gebrauch der gegenwärtig besten externen, wissenschaftlichen Evidenz für Entscheidungen in der medizinischen Versorgung individueller Patienten.

Die Praxis der EBM bedeutet die Integration individueller klinischer Expertise mit der bestmöglichen externen Evidenz aus systematischer Forschung.

Externe Evidenz: Ergebnisse patientenorientierter Forschung

Interne Evidenz:

- Erfahrenes (Beobachten, Messen, Wiedererkennen)
- muss geübt und gebahnt werden (u.a. durch ext. Evidenz, pathophysiologisches Wissen & Lernen)

Absolute und relative Risikoreduktion, NNT

1. Unter Standardtherapie erleiden 20% ein unerwünschtes Ereignis.
2. Unter der experimentellen Therapie erleiden 15% das unerwünschte Ereignis

Risiko bei Standardtherapie	Risiko bei experimenteller Therapie	Relative Risikoreduktion (RRR)	Absolute Risikoreduktion (ARR)	Number Needed to Treat (NNT)
R_S	R_E	$\frac{R_S - R_E}{R_S}$	$R_S - R_E$	$\frac{1}{ARR}$
0.20	0.15	$\frac{0.20 - 0.15}{0.20} = 0.25$	$0.20 - 0.15 = 0.05$	$1/0.05 = 20$

Pro 20 Patienten, die mit der exp. Therapie behandelt werden, wird ein unerwünschtes Ereignis im Vergleich zur Standardtherapie verhindert

Klinische Informationssysteme

korrekte Patientenidentifikation ist wichtig!

Fehler bei Wiederaufnahme:

- **falsch-negative** Identifikation: z.B. Patient sagt, er war noch nie da -> neue Akte wird angelegt
- **falsch-positive** Identifikation: falsche Zuordnung z.B. aufgrund des gleichen Namens und Geburtstag oder die Frau an der Anmeldung hat die Chipkarte durchgezogen als noch das Fenster vom vorherigen Patienten offen war -> Überschreibung.
z.B. Herr Müller hat die Krankengeschichte von Herrn Kunz

Formulierungen(Beispiele RR, RD, RRR, Hazard Ratio, KI, Hypothesen)

- RR (relatives Risiko):** Das Hypertonie-Risiko ist in der Nifedipin-Gruppe 0,41 mal so hoch wie in der Placebogruppe, d.h. das Hypertonie-Risiko wird durch die Gabe des Medikaments im Vergleich zur Placebogruppe um relativ 59% erniedrigt.
- RD (Risikodifferenz):** Das Hypertonierisiko in der Nifedipin-Gruppe ist um 20 Prozentpunkte niedriger als in der Placebogruppe, d.h. das Grundrisiko einer Hypertonie in der Schwangerschaft wird durch das Medikament um 20 Prozentpunkte (20%) gesenkt.
- RRR (relative Risikoreduktion):** Das relative Risiko einer Blutung für Patienten mit der Aspirintherapie beträgt im Vergleich zu den Patienten mit einer oralen Antikoagulation 3.8%/11.2% = 0.34 oder 34%. In der Aspirin-Gruppe traten demnach $1 - 0.34 = 0.66$ oder 66% weniger Blutungen auf. Dieser Wert sagt aber nichts darüber aus, wie viele Ereignisse wirklich durch eine Behandlung verhindert werden, sondern nur wie viele Ereignisse in der Behandlungsgruppe *im Vergleich, resp. relativ* zur Vergleichsgruppe seltener sind.
- Hazard Ratio:**
(4,16) Das Risiko für Patienten in der Placebogruppe ist 4 mal so hoch eine Progression zu erleiden im Vergleich zur Gruppe, die mit Medikamenten behandelt wurde.

Konfidenzintervall: (und NNT)	Die unbekannte Anzahl der zu behandelnden Patienten die mit OP1 statt OP2 behandelt werden, damit ein Patient weniger postoperative Schmerzen angibt, liegt mit 95%iger Wahrscheinlichkeit zwischen 3 und 13 Personen.
KI inkl. Nulllage: (RR)	Das Relative-Risiko-Konfidenzintervall von [0,69;6,24] würde bedeuten, dass nicht mit ausreichender Sicherheit ausgeschlossen werden kann, dass kein Effekt besteht (RR = 1) oder sogar das Placebo überlegen ist.
Nullhypothese H_0 :	Die durch die beiden Sedativa hervorgerufenen mittleren Veränderungen der Schlafdauer unterscheiden sich im Mittel nicht.
Arbeitshypothese H_A :	Der Unterschied in der mittleren Veränderung der Schlafdauer der Patienten wird durch die beiden Sedativa hervorgerufen und ist nicht zufällig.