

| | | | |
|--------------------------|----------------|----------------|-----------|
| AWMF-Register Nr. | 041/004 | Klasse: | S3 |
|--------------------------|----------------|----------------|-----------|

Glossar

(modifiziert und erweitert nach: Das Deutsche Cochrane Zentrum: Glossar www.cochrane.de/de/cochrane-glossar sowie www.awanet.ch/nv/EBM/Glossar.asp)

Beobachtungsstudien

Querschnittsstudien (http://de.wikipedia.org/wiki/Querschnitt_%28empirische_Forschung%29) (engl. cross sectional study) ermitteln eine Momentaufnahme der untersuchten epidemiologischen Daten. Durch den zeitlichen „Schnappschuss“ der epidemiologischen Daten sind die aus der Studie gezogenen kausalen Zusammenhänge zwischen Exposition und Erkrankung schwach und dienen mehr der Generierung von Hypothesen als deren Verifizierung.

Längsschnittstudien (<http://de.wikipedia.org/wiki/L%C3%A4ngsschnittstudie>) (engl. longitudinal study) sind ein Überbegriff für Studien, die regelmäßig Daten der Studienpopulation über einen längeren Zeitraum hinweg erheben. Sie entsprechen periodisch durchgeführten Querschnittsstudien.

Kohortenstudien (http://de.wikipedia.org/wiki/Kohorte_%28Sozialwissenschaft%29) (engl. cohort studies) untersuchen definierte Gruppen von Menschen mit und ohne Exposition einem Risikofaktor gegenüber über eine längere Zeit und messen am Ende des Beobachtungszeitraums den Erkrankungsstatus. Aus der Anzahl Erkrankter unter den Exponierten dividiert durch die Gesamtzahl an Exponierten kann das Risiko der Exponierten für diese Erkrankung gemessen werden. Analog verfährt man für die Nicht-Exponierten. Das Verhältnis des Risikos der Exponierten zum Risiko der Nicht-Exponierten ist das Risikoverhältnis (auch genannt relatives Risiko (http://de.wikipedia.org/wiki/Relatives_Risiko) oder engl. risk ratio) und gibt an, wie stark die Exposition das Risiko der Erkrankung erhöht. Beispielsweise erhöht Rauchen von täglich 20 Zigaretten gegenüber Nicht-Rauchen das Risiko, an Lungenkrebs zu erkranken, um den Faktor 15. Bei prospektiven (http://de.wikipedia.org/wiki/Prospektive_Studie) Kohortenstudien liegen Studienbeginn und Beginn des Beobachtungszeitraums eng beieinander, die Kohorte wird „in die Zukunft“ (prospektiv) verfolgt und der Krankheitsstatus ist noch unbekannt. Retrospektive (http://de.wikipedia.org/wiki/Retrospektive_Studie) Kohortenstudien betrachten bereits vergangene Kohorten, hier sind die Beobachtungen bereits abgeschlossen und der Krankheitsstatus ist bereits bekannt. Sie sind einfacher und kostengünstiger durchzuführen als prospektive Kohortenstudien, allerdings auch anfälliger für Verzerrer (engl. bias), speziell bei der Rekrutierung der Studienteilnehmer, die ja in der Vergangenheit lag und nicht mehr zu beeinflussen ist. Beispiele für Kohortenstudien wäre die Untersuchung von Lungenkrebs bei Asbestarbeitern (<http://de.wikipedia.org/wiki/Asbest>) (exponierte Gruppe) einer Firma und deren Büroangestellten (nicht-exponierte Gruppe).

| | Anzahl Erkrankter | Anzahl Gesunder |
|--------------------------|-------------------|-----------------|
| Anzahl Exponierter | a | b |
| Anzahl Nicht-Exponierter | c | d |

$$\text{Risikoverhaeltnis} = \frac{\frac{a}{a+b}}{\frac{c}{c+d}}$$

Fall-Kontrollstudien (<http://de.wikipedia.org/wiki/Fall-Kontroll-Studie>) (engl. case control study) gehen methodisch den umgekehrten Weg einer Kohortenstudie. Bei einer Fall-Kontrollstudie ist der Krankheitsstatus bekannt und die Exposition unbekannt. Sie eignet sich insbesondere für seltene Erkrankungen, da eine Kohortenstudie sehr viele Teilnehmer haben müsste, um eine statistisch ausreichende Anzahl Erkrankter zu erreichen. Die Studienpopulation der Fall-Kontrollstudie besteht aus Erkrankten und Gesunden, wobei aus statistischen Gründen auf einen Erkrankten auch zwei oder mehr Gesunde kommen können (1:2 matching, 1:n matching). Erst nach der Zuordnung zu den beiden Gruppen wird die Exposition erfasst, um Beeinflussungen des Ergebnisses durch die Beobachter auszuschließen. Ausgewertet wird die Chance (engl. odds) (<http://de.wikipedia.org/wiki/Odds>) der Erkrankten, exponiert zu sein. Sie ergibt sich aus der Zahl der Erkrankten mit Exposition dividiert durch die Zahl der Erkrankten ohne Exposition ('nicht' die Gesamtzahl der Erkrankten). Analog wird die Chance der Gesunden berechnet, exponiert zu sein. Die Division der Chance der Erkrankten durch die Chance der Gesunden ergibt das Chancenverhältnis (engl. odds ratio) (http://de.wikipedia.org/wiki/Odds_ratio). Es entspricht dem Faktor, um den sich die Chance erhöht, durch die Exposition zu erkranken. In einer Fall-Kontrollstudie muss man das Chancenverhältnis und nicht etwa das Risikoverhältnis berechnen, da man durch die willkürliche Wahl der Anzahl von Kontrollpersonen den Nenner des Risikoterms (der Summe von a+b) verzerren würde. Hingegen würde sich eine Verdoppelung von Kontrollen beim Chancenverhältnis wieder rechnerisch herauskürzen (doppelt so viele im Zähler wie im Nenner). Bei seltenen Erkrankungen entspricht das Chancenverhältnis dem Risikoverhältnis. Fall-Kontrollstudien sind grundsätzlich retrospektiv.

| | Anzahl Erkrankter | Anzahl Kontrollen |
|--------------------------|-------------------|-------------------|
| Anzahl Exponierter | a | b |
| Anzahl Nicht-Exponierter | c | d |

$$\text{Odds Ratio} = \frac{a * d}{c * b} = \frac{\frac{a}{c}}{\frac{b}{d}}$$

Bias (Verzerrung oder systematischer Fehler)

Tendenz der Studienergebnisse, systematisch von den „wahren“ Ergebnissen abzuweichen. Systematische Fehler führen entweder zu einer Über- oder Unterschätzung der wahren Wirkung einer Maßnahme. Die Ursachen dafür liegen vor allem im Design und der Durchführung der Studie und führen zu systematischen Unterschieden zwischen den Vergleichsgruppen, z.B. durch die Auswahl der Teilnehmer (Selektionsbias), der Erhebung der Endpunkte (*measurement bias* oder

| |
|---|
| <p>Messungsbias) oder dem Verlust von Teilnehmern in der Studie (<i>attrition bias</i> oder Verschleiß-Bias). Ergebnisse aus Studien mit geringem Risiko für Verzerrungen werden als valide (siehe Validität) angesehen. Weitere mögliche systematische Fehler sind Verzerrungen durch kleine Studien oder die selektive Publikation von positiven (und fehlende Publikation negativer) Studienergebnisse in wissenschaftlichen Zeitschriften (Publikationsbias).</p> |
| <p>Cochrane Collaboration Die Cochrane Collaboration ist ein weltweites Netz von Wissenschaftlern und Ärzten. Ziel ist, systematische Übersichten zur Bewertung von medizinischen Therapien zu erstellen, aktuell zu halten und zu verbreiten. Die Cochrane Collaboration ist eine gemeinnützige Organisation. Die Mitarbeit ist freiwillig. Die Cochrane-Zentren werden von Gesundheitsbehörden, Universitäten oder von wissenschaftlichen Fonds unterstützt. Die Cochrane Collaboration hat ein kostenlos erhältliches Computerprogramm zur Erstellung von Metaanalysen entwickelt (Software RevMan), welche für die Metaanalysen der Leitlinie verwendet wird.</p> |
| <p>Diskrete Variablen Diskrete Variablen weisen im Gegensatz zu kontinuierlichen Variablen nur eine begrenzte Zahl eindeutig voneinander abgrenzbarer Zustände auf (z.B. Augenfarbe: blau, grau, braun, grün). Eine Sonderform sind dichotome Variablen, die lediglich zwei Alternativen aufweisen, z. B. Raucher / Nichtraucher, lebend oder tot, Test-positiv oder Test-negativ. Für die Leitlinie wurde die dichotome Variable „Studienabbruch ja oder nein“ analysiert.</p> |
| <p>Effektmaß Maßzahl, um die Stärke eines Effekts zu quantifizieren. Gebräuchliche Effektmaße in Metaanalysen für dichotome Endpunkte sind das relative Risiko (RR) oder die Odds Ratio (OR), gebräuchliche Effektmaße für kontinuierliche Endpunkte sind die standardisierte mittlere Differenz (SMD) oder die gewichtete mittlere Differenz (weighted mean difference, WMD).</p> |
| <p>Evidenz (evidence) Der Begriff „Evidenz“ im Kontext der Evidenzbasierten Medizin leitet sich vom englischen Wort „evidence“ = Nach-, Beweis ab und bezieht sich auf die Informationen aus klinischen Studien, die einen Sachverhalt erhärten oder widerlegen.</p> |
| <p>Evidenz-basierte Medizin (Evidence-Based Medicine, EbM) Unter Evidenz-basierter Medizin (EbM) oder evidenzbasierter Praxis im engeren Sinne versteht man eine Vorgehensweise des medizinischen Handelns, individuelle Patienten auf der Basis der besten zur Verfügung stehenden Daten zu versorgen. Diese Technik umfasst die systematische Suche nach der relevanten Evidenz in der medizinischen Literatur für ein konkretes klinisches Problem, die kritische Beurteilung der Validität der Evidenz nach klinisch-epidemiologischen Gesichtspunkten; die Bewertung der Größe des beobachteten Effekts sowie die Anwendung dieser Evidenz auf den konkreten Patienten mit Hilfe der klinischen Erfahrung und der Vorstellungen der Patienten.</p> |
| <p>Forest-Plot Die graphische Darstellung der Einzelergebnisse aller eingeschlossenen Studien einer systematischen Übersicht, zusammen mit dem Ergebnis der kombinierten Analyse, falls eine Meta-analyse durchgeführt wurde. Für Cochrane Reviews werden diese Darstellungen standardisiert unter Verwendung des Review-Managers hergestellt.</p> |
| <p>Heterogenität / Homogenität In systematischen Reviews oder Meta-Analysen bezeichnet Homogenität</p> |

(Heterogenität), inwieweit die in den eingeschlossenen Studien gefundenen Effekte ähnlich (homogen) oder verschieden (heterogen) sind. Mit statistischen Heterogenitätstests kann festgestellt werden, ob die Unterschiede zwischen den Studien größer sind, als zufallsbedingt zu erwarten wäre. Als Ursachen für Heterogenität kommen Unterschiede in den Patientencharakteristika, Interventionen oder Endpunkte zwischen den Studien in Frage.

Als statistische Maß für die Heterogenität eines **Effektmaßes** wird in der Leitlinie die I^2 -Statistik verwendet. $I^2 < 25\%$ weisen auf eine geringe Heterogenität (bzw. hohe Homogenität der Studienergebnisse) hin, d.h. alle Studien kamen zu sehr ähnlichen Ergebnissen bzgl. der Wirksamkeit einer Therapie. I^2 zwischen 25-50% weisen auf eine mäßige und $>50\%$ auf eine hohe Heterogenität hin. Die Durchführung einer Meta-Analyse aus heterogenen Studien ($I^2 > 50\%$) ist methodisch problematisch.

Intention-to-treat-Analyse

Analyse-Technik, bei der die Patienten nach ihrer ursprünglichen Gruppenzuteilung analysiert werden, unabhängig davon, ob sie die zugeordnete (intendierte) Therapieform vollständig, partiell oder gar nicht erhalten haben.

Interventionsstudie

Interventionsstudien (engl. intervention studies) verfolgen ähnlich einer prospektiven Kohortenstudie eine Population entlang der Zeit, wobei man den Einfluss einer spezifischen Intervention, meist eine neue Behandlung oder ein neues Medikament, auf das Krankheitsrisiko messen möchte. Vor der Studie wird die Population in den Interventionszweig und den Kontrollzweig geteilt. Während der Studie wird dann aktiv diese Intervention (z.B. Medikament) gegeben, während die Kontrollpopulation unbehandelt bleibt, bzw. eine nicht-wirksame Behandlung bekommt (z.B.: Placebo) (<http://de.wikipedia.org/wiki/Placebo>). Die Auswertung erfolgt ähnlich einer Fall-Kontrollstudie über Chancenverhältnisse (engl. odds ratios). Die Zuordnung zur Behandlungsgruppe und Kontrollgruppe ist der kritische Punkt einer Interventionsstudie, da sich die Teilnehmer in ihren Gesundheitsparametern unterscheiden und man nur den Einfluss der Intervention und nicht dieser Parameter messen möchte. Erfolgt diese Auswahl zufällig und damit nicht gerichtet, spricht man von einer randomisierten, kontrollierten Studie (engl. randomised controlled trial). Diese Studien haben eine besonders starke Kausalität in Bezug auf Intervention und Krankheitsstatus und werden daher in der Medikamententestung eingesetzt.

Klinische Studie

Unschärf definierter Begriff für eine Studie, in der eine Intervention an einer Gruppe von Patienten untersucht wird. Oberbegriff für unterschiedliche Studientypen, z. B. nicht kontrollierte, kontrollierte und randomisierte klinische Studien.

Kohortenstudie

Vergleichende Beobachtungsstudie, in der Personen (Kohorte) mit bzw. ohne eine Intervention / Exposition (zu der sie nicht von dem Studienarzt zugeteilt wurden) über einen definierten Zeitraum beobachtet werden, um Unterschiede im Auftreten der Zielerkrankung festzustellen. Kohortenstudien können prospektiv oder retrospektiv durchgeführt werden.

Komorbidität

Weitere Erkrankungen, welche zusätzlich bei den Personen mit der untersuchten Erkrankung vorliegen (z. B. depressive Störungen beim FMS)

Konfidenzintervall (Vertrauensbereich, *confidence interval* – CI)

Bereich, in dem der „wahre“ Wert einer Messung (**Effektgröße**) mit einer bestimmten Wahrscheinlichkeit erwartet werden kann (üblicherweise 95% Konfidenzintervall). Das wahre Ergebnis der Studie liegt mit einer 95% Wahrscheinlichkeit in diesem Bereich. Die Effektgröße kann dabei z.B. ein

Therapieeffekt oder ein Risiko sein. Das Konfidenzintervall beschreibt die Unsicherheit über die Zuverlässigkeit der
Das Konfidenzintervall einer Effektgröße macht auch eine Aussage über das **Signifikanzniveau** eines Effektes. Wenn das 95% Konfidenzintervall beim Vergleich von zwei Stichprobenmittelwerten (z. B. Vergleich von 2 verschiedenen Therapieverfahren) 0 beinhaltet, kann davon ausgegangen werden, dass sich die Mittelwerte mit einer Irrtumswahrscheinlichkeit von 0.05 nicht signifikant unterscheiden.

Kontinuierliche / stetige Variablen

Im Gegensatz zu **diskreten Variablen** können kontinuierliche Variablen theoretisch eine unendlich große Zahl von Werten entlang eines Kontinuums annehmen. Körpergröße, Gewicht und viele Laborwerte sind kontinuierliche Variablen. In der Leitlinie werden die kontinuierlichen Variablen der Studienendpunkte Schmerzintensität, Müdigkeit, Schlafstörung und gesundheitsbezogene Lebensqualität, welche mit Fragebögen in den Studien erfasst wurden, analysiert.

Meta-Analyse

Statistisches Verfahren, um die Ergebnisse aller Studien, die die gleiche Frage bearbeiten, quantitativ zu einem Gesamtergebnis zusammenzufassen und dadurch die Aussagekraft (Genauigkeit der Effektschätzer) gegenüber Einzelstudien zu erhöhen.

Nullhypothese

Bei der Durchführung statistischer Signifikanztests wird üblicherweise die Hypothese aufgestellt, dass zwischen den verschiedenen Gruppen einer Studie kein Unterschied (Nullhypothese) besteht. Aus statistischer Sicht ist eine Maßnahme wirksam, wenn man durch den statistischen Test die Nullhypothese, dass es zwischen den Ergebnissen der experimentellen und der Kontrollgruppe keinen Unterschied gibt, verwerfen kann

p-Wert

p-Werte (p von *probability*) beschreiben die Wahrscheinlichkeit, dass der beobachtete (oder ein noch extremerer) Effekt einer Studie aufgetreten sein könnte, wenn die **Nullhypothese** richtig und der Effekt auf das Spiel des Zufalls zurückzuführen ist. Je kleiner der Wert, desto deutlicher spricht das beobachtete Ergebnis gegen die Nullhypothese. Es ist eine Konvention, dass ein p-Wert gleich oder kleiner 0.05 als statistisch signifikant angesehen wird. Wenn die Signifikanz von Effekten interpretiert wird, sollten p-Werte immer im Zusammenhang mit **Konfidenzintervallen** verwendet werden.

Placebo

In einer pragmatischen Definition ein meist zu Studienzwecken eingesetztes "Schein-Medikament" ohne pharmakologisch aktive Substanz. Das Placebo darf hinsichtlich seiner äußeren Eigenschaften nicht von der aktiven Behandlung (dem Verumpräparat) unterscheidbar sein, wenn es seinen Zweck erfüllen soll.

Random Effects Model

Statistisches Modell zur Berechnung zusammengefasster (gepoolter) Ergebnisse, bei denen im Gegensatz zum Fixed Effect Modell Effektunterschiede zwischen verschiedenen Studien berücksichtigt werden. In die Genauigkeit der Schätzung des gemeinsamen Effekts geht daher nicht nur die Variation ein, die innerhalb der Studien beobachtet wird, sondern auch die Variation zwischen den Studien. Ein Beispiel ist die Methode nach DerSimonian & Laird.

Das Random Effects Model führt im Gegensatz zum Fixed Model, das einen gemeinsamen Effekt der Behandlung in den Studien annimmt, zu einer konservativeren Schätzung (in der Regel geringeren) Effektstärken.

Mit verschiedenen Verfahren kann gewichtet werden, zu welchem Anteil jede

einzelne Studie in das zusammengefasste Ergebnis der Meta-analyse eingeht. In den Meta-Analysen der Leitlinie wird die Methode der inversen Varianz benutzt, d.h. Studien mit großen Teilnehmerzahlen und einer größeren Präzision (erfasst mittels der Methode der inversen Varianz) erhalten ein größeres Gewicht bei dem zusammengefasste Ergebnis der Meta-analyse als Studien mit kleinen Teilnehmerzahlen bzw. großer Schwankungsbreite der Ergebnisse.

Randomisierung

Verfahren, das eine zufällige Verteilung der Patienten auf eine Therapie- und eine Kontrollgruppe bewirkt (**s. randomisierte kontrollierte Studie**). Dies kann durch (computergenerierte) Zufallszahlen oder andere Mechanismen erreicht werden. Damit soll sichergestellt werden, dass alle Teilnehmer die gleiche Chance haben, der einen oder anderen Gruppe zugeordnet zu werden und es wahrscheinlich ist, dass sich (bei ausreichender Studiengröße) bekannte wie unbekannte Risiko- und Prognosefaktoren ausgeglichen auf die beiden Gruppen verteilen. Wenn sich zwischen den beiden Gruppen in den Endpunkten ein Unterschied zeigt, kann dieser tatsächlich der experimentellen Intervention zugeordnet werden. Die Randomisierung ist das entscheidende Instrument zur Verhinderung einer Selektionsbias und damit eines der wichtigsten Mittel zur Sicherung der internen Validität einer klinischen kontrollierten Studie.

Randomisierte kontrollierte Studie (RCT)

Eine experimentelle Studie, bei der die Patienten nach einem Zufallsverfahren (**Randomisierung**) und mit **verdeckter Zuordnung** auf die Therapie- bzw. die Kontrollgruppe verteilt (randomisiert) und auf das Auftreten der festgelegten Endpunkte in den einzelnen Gruppen nachbeobachtet werden.

Relatives Risiko (RR)

Effektmaß für **dichotome Variablen**. Das relative Risiko in einer Therapiestudie bezeichnet das Verhältnis zwischen dem Risiko in der experimentellen Gruppe und dem Risiko in der Kontrollgruppe. Ein relatives Risiko von 1 bedeutet, dass zwischen den Vergleichsgruppen kein Unterschied besteht. Bei ungünstigen Ereignissen (z. B. Abbruchrate wegen Nebenwirkungen) zeigt ein $RR < 1$, dass die experimentelle Intervention wirksam ist, um das Auftreten von ungünstigen Ereignissen zu senken.

Revman, Review-Manager

Software der Cochrane Collaboration zur Erstellung von Reviews und Meta-Analysen (<http://www.cc-ims.net/>)

Standardabweichung

Maß für die Streuung von Messwerten um den Durchschnittswert.

Standardisierte mittlere Differenz

Die standardisierte mittlere Differenz (SMD) ist ein statistische Maß für die Größe eines Unterschiedes zwischen zwei unabhängigen Stichproben (z. B. Vergleich eines Medikamentes mit Placebo). Es gibt verschiedene statistische Berechnungsmethoden für standardisierte mittlere Differenz, z. B. Cohen's d oder Hedges' g. Die für die Meta-Analysen verwendete Software **Revman 5.0** verwendet Hedges'g als Maß für die standardisierte mittlere Differenz (Differenz der Mittelwerte der beide Stichproben dividiert durch die gepoolte Standardabweichung der Ergebnisse der Studienteilnehmer mit einem Korrekturfaktor für kleine Stichprobenumfänge).

SMDs haben die wichtige Eigenschaft, dass ihr Wert unabhängig von den verwendeten Skalen ist, d.h. es können Maße mit unterschiedlichen Skalierungen verwendet werden (z. B. können Ergebnisse einer 0-10 und einer 1-5 Schmerzskala kombiniert werden). SMD messen die Zahl der Standardabweichungen zwischen den Mittelwerten. SMDs sind jedoch schwierig zu interpretieren:

Interpretation von standardisierten Mittelwertdifferenzen

| Standardisierte Mittelwertdifferenz | Der Wert (z. B. Schmerzreduktion) der Durchschnittsperson in der Experimentalgruppe ist größer als der Wert von .. % der Personen der Kontrollgruppe |
|-------------------------------------|--|
| 0.0 | 50% |
| 0.1 | 54% |
| 0.2 | 58% |
| 0.3 | 62% |
| 0.4 | 66% |
| 0.5 | 69% |
| 0.6 | 73% |
| 0.7 | 76% |
| 0.8 | 79% |
| 0.9 | 82% |
| 1.0 | 84% |
| 1.2 | 88% |
| 1.4 | 92% |
| 1.6 | 95% |
| 1.8 | 96% |
| 2.0 | 98% |
| 2.5 | 99% |
| 3.0 | 99.9% |

Eine Effektstärke von 0.8 bedeutet, dass der Wert der Durchschnittsperson in der Experimentalgruppe den Wert von 79% der Personen der Kontrollgruppe übersteigt.

Coe, R. (2002). It's the Effect Size, Stupid. What Effect Size is and Why it is Important. Paper presented at the Annual Conference of the British Educational Research Association, University of Exeter, England, 12-14 September 2002.

Häufig wird die Größe einer standardisierten Mittelwertdifferenz nach den von Cohen vorgeschlagenen Kategorien beurteilt.

SMD <0.2: Nicht bedeutsam

SMD 0.2-0.5: Gering

SMD 0.5-0.8: Mäßig

SMD >0.8: Hoch

Statistischer Signifikanztest

Ziel eines **statistischen Signifikanztests** ist die Ablehnung (Verwerfung) der Nullhypothese. Falls diese nicht verworfen werden kann (z. B. weil nicht genügend Beobachtungen vorhanden sind), besteht aus statistischer Sicht allerdings kein Grund, von der Gültigkeit der Nullhypothese auszugehen, d.h. der fehlende Nachweis der Wirksamkeit einer Therapie ist kein Beweis für eine fehlende Wirksamkeit. Ein **statistischer Test** kann also lediglich zu einer Annahme der Alternativhypothese, nicht aber zu einer Annahme der Nullhypothese führen.

Statistische Signifikanz

Ein statistisch signifikantes Ergebnis einer Studie ist ein Ergebnis, das gegen die **Nullhypothese** spricht. Die Aussage basiert auf einem statistischen Test, der zur Prüfung einer vorab festgelegten Hypothese mit vorab festgelegter Irrtumswahrscheinlichkeit durchgeführt wird. Statistische Signifikanz darf nicht mit klinischer Relevanz gleich gesetzt werden, z. B. kann der Unterschied zwischen der Experimental- und der Kontrollgruppe bei großen Teilnehmerzahlen signifikant sein, der Unterschied (z. B. 5 Einheiten auf einer 0-100) Schmerzskala jedoch für den Patienten bzw. Behandler nicht bedeutsam sein.

Systematische Übersicht (engl. Systematic review)

Sekundärforschung, bei der zu einer klar formulierten Frage alle verfügbaren Primärstudien systematisch und nach expliziten Methoden identifiziert, ausgewählt und kritisch bewertet und die Ergebnisse extrahiert und deskriptiv oder mit statistischen Methoden quantitativ (**Meta-Analyse**) zusammengefasst werden.

Validität (Validität, Glaubwürdigkeit)

Die Validität bezeichnet das Ausmaß, mit dem die Ergebnisse einer Studie die „wahren“ Effekt einer Intervention / Exposition wiedergegeben werden, d.h. frei von systematischen Fehlern (**Bias**) sind. Die innere Validität beruht auf der methodischen Qualität des Studiendesigns.

In der Leitlinie wird die methodische Qualität (interne Validität) der Studien an Hand der Angemessenheit der **Randomisierung**, der **verdeckten Zuordnung** der Behandlung, der **Verblindung** der Endpunkt-Auswerter und durch die Durchführung einer **Intention-to-treat-Analyse** bewertet.

Die externe Validität der Studien beschreibt die Übertragbarkeit von Studienergebnissen auf die Patienten in der Routineversorgung, d.h. auf Patienten, die nicht an der Studie teilgenommen haben. Die häufigste Einschränkung der externen Validität liegt in der Auswahl der Personen, die an der Studien teilnehmen, d.h. ob die Studienteilnehmer repräsentativ (z. B. Alter, Geschlecht, soziale Schicht, Krankheitsdauer und – schwere sowie Begleiterkrankungen) sind für die Patienten, die in der klinischen Routine gesehen werden. Eine weitere mögliche Einschränkung der externen Validität sind die Rahmenbedingungen der Studie, z. B. ob die Dosis

oder der Umfang der Therapie (z. B. einer psychologischen Behandlung) der Dosis oder des Umfangs einer Therapie entspricht, wie sie in der klinischen Routine angeboten werden kann.

In der Leitlinie wird die externe Validität der Studien an Hand der Ein- und Ausschlusskriterien der Studienteilnahme überprüft. Es wird überprüft, ob Patienten mit häufigen Komorbiditäten des FMS, nämlich entzündlich-rheumatische Erkrankungen und Angst- bzw. Depressive Störungen in die Studie aufgenommen wurden.

Verblindung (Blinding)

Geheimhaltung der Gruppenzuordnung (Therapie oder Kontrolle) vor Patienten, Studienärzten, Pflegepersonal oder Auswertern, die an einer Studie teilnehmen. Damit soll verhindert werden, dass durch das Wissen um die Gruppenzugehörigkeit die Therapieantwort der Patienten, das Verhalten der Ärzte oder die Bewertung der Ergebnisse beeinflusst wird. In einfach-blinden Studien wissen nur die Patienten nicht über ihre Zuordnung Bescheid, in doppel-blinden Studien bleibt die Zuordnung Patient und behandelndem Arzt verborgen. Die Verblindung von Ärzten und Patienten ist nicht immer durchführbar (z. B. beim Vergleich von chirurgischen mit medikamentösen Verfahren), wobei eine Verblindung der Endpunkt-Auswerter in der Regel möglich ist.

Verdeckte Zuordnung (*concealment of allocation*)

Methodisches Verfahren zum Schutz vor Selektionsbias. Geheimhaltung der randomisierten Zuteilungsfolge zu Therapie- oder Kontrollgruppe bis zum Zeitpunkt des Studieneinschlusses und der Zuordnung des Patienten zu einer Studiengruppe.