

PEDro-skala – Deutsch

1. Die Ein- und Ausschlusskriterien wurden spezifiziert nein ja wo:
 2. Die Probanden wurden den Gruppen randomisiert zugeordnet (im Falle von Crossover Studien wurde die Abfolge der Behandlungen den Probanden randomisiert zugeordnet) nein ja wo:
 3. Die Zuordnung zu den Gruppen erfolgte verborgen nein ja wo:
 4. Zu Beginn der Studie waren die Gruppen bzgl. der wichtigsten prognostischen Indikatoren einander ähnlich nein ja wo:
 5. Alle Probanden waren geblindet nein ja wo:
 6. Alle Therapeuten/Innen, die eine Therapie durchgeführt haben, waren geblindet nein ja wo:
 7. Alle Untersucher, die zumindest ein zentrales Outcome gemessen haben, waren geblindet nein ja wo:
 8. Von mehr als 85% der ursprünglich den Gruppen zugeordneten Probanden wurde zumindest ein zentrales Outcome gemessen nein ja wo:
 9. Alle Probanden, für die Ergebnismessungen zur Verfügung standen, haben die Behandlung oder Kontrollanwendung bekommen wie zugeordnet oder es wurden, wenn dies nicht der Fall war, Daten für zumindest ein zentrales Outcome durch eine ‚intention to treat‘ Methode analysiert nein ja wo:
 10. Für mindestens ein zentrales Outcome wurden die Ergebnisse statistischer Gruppenvergleiche berichtet nein ja wo:
 11. Die Studie berichtet sowohl Punkt- als auch Streuungsmaße für zumindest ein zentrales Outcome nein ja wo:
-

Die PEDro-Skala basiert auf der Delphi Liste, die von Verhagen und Kollegen an der Universität von Maastricht, Abteilung für Epidemiologie, entwickelt wurde (Verhagen AP et al (1998). The Delphi list: a criteria list for quality assessment of randomised clinical trials for conducting systematic reviews developed by Delphi consensus. Journal of Clinical Epidemiology, 51(12):1235-41). Diese Liste basiert auf einem „Expertenkonsens“, und größtenteils nicht auf empirischen Daten. Zwei zusätzliche Items, die nicht Teil der Delphi Liste waren, wurden in die PEDro-Skala aufgenommen (Kriterien 8 und 10). Wenn mehr empirische Daten zur Verfügung stehen, könnte es in Zukunft möglich werden, die einzelnen Items zu gewichten, so dass eine PEDro-Punktzahl die Bedeutung individueller Items widerspiegelt.

Der Zweck der PEDro-Skala ist es, Benutzern der PEDro-Datenbank dabei zu helfen, schnell festzustellen, welche der tatsächlich oder vermeintlich randomisierten kontrollierten Studien (d.h. RCTs oder CCTs), die in der PEDro-Datenbank archiviert sind, wahrscheinlich intern valide sind (Kriterien 2-9) und ausreichend statistische Information beinhalten, um ihre Ergebnisse interpretierbar zu machen (Kriterien 10-11). Ein weiteres Item (Kriterium 1), welches sich auf die externe Validität (Verallgemeinerungsfähigkeit von Ergebnissen) bezieht, wurde übernommen, um die Vollständigkeit der Delphi Liste zu gewährleisten. Dieses Kriterium wird jedoch nicht verwendet, um die PEDro-Punktzahl zu berechnen, die auf der PEDro Internetseite dargestellt wird.

Die PEDro-Skala sollte nicht als Maß für die „Validität“ der Schlussfolgerungen einer Studie verwendet werden. Insbesondere warnen wir Benutzer der PEDro-Skala, dass Studien, die einen signifikanten Behandlungseffekt anzeigen, und die hohe Punktzahlen auf der PEDro-Skala erreichen, nicht notwendigerweise den Nachweis dafür erbringen, dass die entsprechenden Behandlungen klinisch sinnvoll sind. Weiterführende Überlegungen beinhalten, ob der Behandlungseffekt groß genug gewesen ist, um lohnenswert zu sein, ob die positiven Effekte der Behandlung die negativen aufwiegen, und wie das Kosten-Nutzen-Verhältnis der Behandlung ist. Die PEDro-Skala sollte nicht dazu verwendet werden, die „Qualität“ von Studien aus unterschiedlichen therapeutischen Bereichen zu vergleichen, und zwar hauptsächlich deswegen nicht, weil es in manchen Bereichen der physiotherapeutischen Praxis nicht möglich ist, allen Kriterien der Skala gerecht zu werden.

Die PEDro-Skala wurde zuletzt am 21. Juni 1999.

Die deutsche Übersetzung der PEDro-Skala wurde erstellt von Stefan Hegenscheidt, Angela Harth und Erwin Scherfer.
Die deutsche Übersetzung wurde im April 2008 fertiggestellt und wurde im Februar 2010 geändert.

Hinweise zur Handhabung der PEDro scale:

- Für alle Kriterien **Punkte werden nur vergeben, wenn ein Kriterium eindeutig erfüllt ist**. Falls beim genauen Lesen einer Arbeit die Möglichkeit besteht, dass ein Kriterium nicht erfüllt wurde, sollte kein Punkt für dieses Kriterium vergeben werden.
- Kriterium 1 Dieses Kriterium gilt als erfüllt, wenn berichtet wird, wie die Probanden rekrutiert wurden, und wenn eine Liste mit Kriterien dargestellt wird, die genutzt wurde, um zu entscheiden, wer geeignet war an der Studie teilzunehmen.
- Kriterium 2 Wenn in einem Artikel steht, dass die Zuordnung zu den Gruppen randomisiert erfolgte, so wird dies von der Studie angenommen. Die genaue Methode der Randomisierung muss dabei nicht näher spezifiziert sein. Methoden wie Münz- oder Würfelwürfe sollten als Randomisierung angesehen werden. Quasi-randomisierte Zuordnungsverfahren wie die Zuordnung durch Krankenaktennummern im Krankenhaus, Geburtsdatum, oder alternierende Zuordnungen, erfüllen dieses Kriterium nicht.
- Kriterium 3 *Verborgene Zuordnung* bedeutet, dass die Person, die entschieden hat ob der jeweilige Proband für eine Teilnahme geeignet war oder nicht, zum Zeitpunkt dieser Entscheidung nicht wissen konnte, welcher Gruppe der jeweilige Proband zugeordnet werden würde. Für dieses Kriterium wird auch dann ein Punkt vergeben, wenn über eine verdeckte Zuordnung nicht berichtet wird, aber in dem Bericht zum Ausdruck kommt, dass die Zuordnung mit Hilfe blickdichter Briefumschläge erfolgte, oder dass die Allokation über Kontaktaufnahme mit einem unabhängigen Verwalter des Allokationsplans, der sich ‚nicht am Ort der Studiendurchführung‘ befand oder ‚nicht anderweitig an der Studie beteiligt‘ war, erfolgte.
- Kriterium 4 In Studien, die therapeutische Interventionen untersuchen, muss jeweils vor Beginn der Intervention mindestens eine Messung hinsichtlich des Schweregrades des zu behandelnden Zustandes, und mindestens ein anderes *zentrales Outcome* beschrieben werden (Eingangsmessungen). Der Gutachter muss ausreichend davon überzeugt sein, dass sich klinisch signifikante Unterschiede in den Gruppen-Outcomes nicht allein schon aufgrund von Unterschieden in den prognostischen Variablen zu Beginn der Studie (also zum Baseline-Zeitpunkt) erwarten ließen. Dieses Kriterium gilt auch dann als erfüllt, wenn nur Baseline-Daten für diejenigen Probanden beschrieben werden, welche bis zum Ende an der Studie teilgenommen haben.
- Kriterien 4,7-11 *Zentrale Outcomes* sind jene Outcomes, welche das primäre Maß für eine Effektivität (oder eine fehlende Effektivität) der Therapie darstellen. In den meisten Studien wird mehr als eine Variable zur Outcome-Messung verwendet.
- Kriterien 5-7 *Blindung* bedeutet, dass die betreffende Person (Proband/In, Therapeut/In oder Untersucher/In) nicht gewusst hat, welcher Gruppe der Proband zugeordnet worden ist. Außerdem wird eine Blindung von Probanden und Therapeuten nur dann als gegeben angenommen, wenn davon ausgegangen werden kann, dass sie nicht in der Lage gewesen wären, zwischen den Behandlungen, die in den verschiedenen Gruppen ausgeführt wurden, zu unterscheiden. In Studien, in denen *zentrale Outcomes* von den Probanden selbst angegeben werden (z.B. Visuelle Analog Skala oder Schmerztagebücher), gilt der Untersucher als geblindet, wenn der Proband geblindet war.
- Kriterium 8 Dieses Kriterium gilt nur dann als erfüllt, wenn die Studie *sowohl* über die Anzahl der ursprünglich den Gruppen zugeordneten Probanden, *als* auch über die Anzahl der Probanden, von denen tatsächlich zentrale Outcomes festgehalten werden konnten, Auskunft gibt. Bei Studien mit Outcome-Messungen zu mehreren Messzeitpunkten, muss mindestens ein *zentrales Outcome* bei mehr als 85% der Probanden zu einem dieser Zeitpunkte gemessen worden sein.
- Kriterium 9 Eine *Intention to treat* Analyse bedeutet, dass in den Fällen, in denen Probanden die zugeordnete Behandlung (oder Kontrollanwendung) nicht erhalten haben und in denen Ergebnismessungen möglich waren, die Messwerte so analysiert werden, als ob die Probanden die zugeordnete Behandlung (oder Kontrollanwendung) erhalten hätten. Wird eine Analyse nach der ‚Intention to treat‘ Methode nicht erwähnt, gilt dieses Kriterium dennoch als erfüllt, falls explizit zum Ausdruck kommt, dass alle Probanden die Behandlungen oder Kontrollanwendungen wie zugeordnet erhalten haben.
- Kriterium 10 Ein *Zwischen-Gruppen-Vergleich* beinhaltet einen statistischen Vergleich einer Gruppe mit einer anderen Gruppe. Abhängig vom jeweiligen Studiendesign kann es sich dabei um den Vergleich von zwei oder mehr verschiedenen Behandlungen, oder auch um den Vergleich einer Behandlung mit einer Kontrollanwendung (z.B. Placebo-Behandlung, Nicht-Behandlung, Scheinbehandlung) handeln. Die Analyse kann als einfacher Vergleich der Outcomes zwischen den Gruppen erfolgen, die nach einer durchgeführten Behandlung gemessen wurden, oder auch als Vergleich der Veränderungen in einer Gruppe mit den Veränderungen in einer anderen Gruppe (wurde eine faktorielle Varianzanalyse durchgeführt, um die Daten zu analysieren, so wird dies im letzteren Fall häufig als eine ‚Gruppe x Zeit Interaktion‘ berichtet). Der Vergleich kann als Hypothesentestung (die einen ‚p‘-Wert liefert, der die Wahrscheinlichkeit dafür angibt, dass der Unterschied zwischen den Gruppen rein zufällig entstanden ist) oder als Schätzung (z.B. der Differenz des Medians oder des arithmetischen Mittels, der Unterschiede in den Prozentanteile, oder der Number Needed to Treat, oder des relativen Risikos oder der ‚Hazard Ratio‘¹) mit einem dazugehörigen Konfidenz-Intervall durchgeführt werden.
- Kriterium 11 Ein *Punktmaß* ist ein Maß der Größe des Behandlungseffekts. Der Behandlungseffekt kann als Differenz in den Outcomes zwischen zwei Gruppen beschrieben werden, oder auch als Outcome in jeder der Gruppen. *Streuungsmaße* können sein: Standardabweichungen, Standardfehler, Konfidenzintervalle, Interquartilsabstände (oder andere Quantilsabstände), und Ranges. Punktmaße und/oder Maße der Streuung können graphisch dargestellt sein (z.B. können Standardabweichungen als Balkendiagramm dargestellt werden), so lange diese Darstellungen eindeutig sind (z.B. so lange klar ist ob die Fehlerbalken Standardabweichungen oder Standardfehler darstellen). Für kategoriale Outcomes (nominal- oder ordinalskaliert) gilt dieses Kriterium als erfüllt, wenn die Anzahl der Probanden für jede Kategorie in jeder Gruppe angegeben ist.

¹ Der Begriff Hazard Ratio (‚Risikoeintrittsquotient‘) wird auch in der deutschen medizinischen Fachliteratur verwendet. Die Hazard Ratio ist der Quotient aus den Eintrittswahrscheinlichkeiten (Ereignisdichten) in den zu vergleichenden Gruppen.