

Evidenzstufen und Empfehlungsgrade: Was Studien wert sind

Je besser das Design einer Studie ist, desto höher ist ihre Qualität. Je höher ihre Qualität ist, desto höher ist ihre Evidenzstufe. Und je höher ihre Evidenzstufe ist, desto eher können Wissenschaftler daraus eine Therapieempfehlung ableiten.

Ob eine Studie aussagekräftig und auf andere Patienten übertragbar ist, hängt vorwiegend von ihrem Design ab. Stehen Experten bei der Erstellung von Leitlinien viele Studien mit gutem Design zur Verfügung, wirkt sich das positiv auf den Grad ihrer Therapieempfehlungen aus.

Design bestimmt Evidenzstufe

Wissenschaftliche Arbeiten, deren Aussagekraft begrenzt ist, sind zum Beispiel Einzelfallstudien (engl.: case studies): Wenn der Autor darin beispielsweise feststellt, dass sich die Hyperemesis der schwangeren Patientin im Verlauf der Akupunktur-Behandlung verbessert, kann er dieses Ergebnis nicht auf alle Schwangeren mit Hyperemesis übertragen. Denn er weiß weder, ob alle Schwangeren mit Hyperemesis von Akupunktur profitieren, noch kann er belegen, dass die Behandlung an sich diese Verbesserungen erbracht hat. Aus Einzelfällen

erfährt man lediglich etwas über die beobachteten und dargestellten Veränderungen, weitere Überlegungen sind nur Mutmaßungen. Einzelfallstudien haben somit eine niedrige Evidenzstufe. Zu den Arbeiten mit hohen Evidenzstufen gehören zum Beispiel randomisierte kontrollierte Studien, da sie eine höherwertige wissenschaftliche Basis bieten, um über die Wirksamkeit einer Therapie urteilen zu können.

Evidenzstufen nicht einheitlich

Je hochwertiger das Design einer Studie also ist, desto höher ist ihre Evidenzstufe. Je nachdem, welche Institution diese Einordnung vornimmt, können sich die einzelnen Evidenzstufen allerdings voneinander unterscheiden.

Die Agency for Health Care Policy and Research (AHCPR) – ein Forschungszweig des amerikanischen Gesundheitsdienstes – verwendet beispielsweise sechs Evidenzgrade (►Tab. 1). Die höchste Aussagefähigkeit haben nach dieser Einteilung Studien der Evidenzstufe Ia. Dazu gehören systematische Reviews randomisierter kontrollierter Studien.

Evidenzklassen beeinflussen Therapieempfehlungen

Wenn Experten Leitlinien veröffentlichen (z. B. die neue S3-Leitlinie der Deutschen Gesellschaft für Gynäkologie und

Geburtshilfe zur Sectio) geben sie darin Empfehlungen, welche Therapieoptionen man bei bestimmten Krankheitsbildern verwenden sollte. Diese Empfehlungen hängen wiederum von den Evidenzstufen der Studien ab, die die Experten für ihre Leitlinien heranziehen können: Je hochwertiger die zum Thema gefundenen Arbeiten sind, desto höher ist der Grad der Therapieempfehlung. Auch dabei gibt es verschiedene Einteilungen. Die Arbeitsgemeinschaft der wissenschaftlichen medizinischen Fachgesellschaften e. V. (AWMF) verwendet drei Grade:

- Eine Empfehlung mit dem Grad A verfügt über schlüssige Literatur von guter Qualität, die mindestens eine randomisierte Studie mit den Evidenzgraden Ia oder Ib enthält.
- Eine Grad-B-Empfehlung wird mit gut durchgeführten, nicht randomisierten klinischen Studien der Evidenzgrade II oder III belegt.
- Eine Empfehlung mit dem Grad C kennzeichnet Berichte und Meinungen von Experten oder klinische Erfahrungen anerkannter Autoritäten, die die Evidenzstufe IV aufweisen. Direkt anwendbare klinische Studien guter Qualität fehlen bei dieser Empfehlung.

Qualität und Empfehlungen einheitlich beurteilen

Die unterschiedlichen Einteilungssysteme erschweren den Vergleich von Studien und Therapieempfehlungen. Aus diesem Grund entwickelten Experten das GRADE-System. Dessen Ziel ist es, Qualität und Empfehlungsstärken einheitlich bewerten und daraus Empfehlungen formulieren zu können. Mehr Informationen dazu gibt es unter www.gradeworkinggroup.org.

Über diesen Artikel

Dieser Artikel basiert auf der Veröffentlichung, die in der Zeitschrift *physiopraxis*, Ausgabe 2 / 2010, erschienen ist: Mehrholz J. Evidenzstufen und Empfehlungsgrade – Was Studien wert sind. *physiopraxis* 2010; 8(6): 20 DOI: 10.1055/s-0030-1262228 → www.thieme-connect.de

►Tab. 1 Evidenzgrade von Studien

Evidenzstufe	Beurteilung
I a	ein systematisches Review auf der Basis methodisch hochwertiger kontrollierter randomisierter Studien (RCTs)
I b	eine ausreichend große, methodisch hochwertige RCT
II a	eine hochwertige Studie ohne Randomisierung, zum Beispiel eine Kohortenstudie
II b	eine hochwertige Studie eines anderen Typs quasi-experimenteller Studien
III	eine methodisch hochwertige nicht experimentelle Studie
IV	Meinungen und Überzeugungen von angesehenen Autoritäten (aus klinischer Erfahrung), Expertenkommissionen, beschreibende Studien, Fallstudien

Tab. Einteilung von Evidenzstufen nach der AHCPR

Autorinnen / Autoren

Prof. Jan Mehrholz
Katharina Kerlen-Petri

Bibliografie

Die Hebamme 2020; 33: 18–19
DOI 10.1055/a-1247-8674
© 2020. Thieme. All rights reserved.
Georg Thieme Verlag KG, Rüdigerstraße
14, 70469 Stuttgart, Germany