

## Leitlinie

### Evidenzbewertung in der Literatur

Leitlinien sind Orientierungshilfen für Ärzte, Patienten und nicht-ärztliche im Gesundheitssystem tätige Personen im Sinne von "Handlungs- und Entscheidungskorridoren", von denen in begründeten Fällen abgewichen werden kann oder sogar muss. Augenärztliche Leitlinien beschreiben, was Augenärzte für eine angemessene Patientenversorgung in der Praxis für geboten halten. Dies entspricht in vielen Fällen nicht dem Leistungsniveau der gesetzlichen Krankenversicherung in Deutschland (siehe [Präambel](#)). Leitlinien sollen die Transparenz erhöhen, in definierten Fragestellungen Wissen aus unterschiedlichen Quellen bündeln und unterschiedliche Standpunkte berücksichtigen und diskutieren [1]. Im Gegensatz zu Richtlinien sind sie rechtlich nicht verbindlich.

#### Evidenzgrade

Für die Vergabe der Evidenzgrade stehen verschiedene Klassifikationen zur Verfügung

Beispiele:

Oxford: [www.cebm.net](http://www.cebm.net)

SIGN: [www.sign.ac.uk](http://www.sign.ac.uk)

GRADE: [www.gradeworkinggroup.org](http://www.gradeworkinggroup.org)

## Oxford

Clinical Evidence: A Practical Approach to Levels of Evidence (2011) - Levels of Evidence (2009)

**Oxford Centre for Evidence-based Medicine - Levels of Evidence (March 2009)**

What are we to do when the best evidence from the best of other clinical studies meets with the irreducible object of flawed evidence? All we can do is our best: give the evidence that best addresses the issue in the situation an OX level 3 or 4 rating. The OXEM Levels of Evidence document sets out one mechanism for operationalising this process for different question types. (For definitions of terms, visit our site: [www.oxfordjournals.org](http://www.oxfordjournals.org))

Level	Therapy / Prevention, Aetiology / Harm	Prognosis	Diagnosis	Differential diagnosis / Symptomatology / Pathophysiology	Etiology and disease analysis
1a	SR (with homogeneity) of RCTs	SR (with homogeneity) of randomised cohort studies, CDR, or meta-analysis of RCTs	SR (with homogeneity) of Level 2 studies or SR (with homogeneity) of RCTs	SR (with homogeneity) of prospective cohort studies	SR (with homogeneity) of Level 1 evidence studies
1b	Individual RCT (with low risk of bias) (with confidence interval)	Individual randomised cohort study with low risk of bias, CDR, published in a single publication	Individual study with low risk of bias, CDR, published in a single publication	Prospective cohort study with good follow-up	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, and including 95% CI and sensitivity analysis
1c	All or none	All or none case series	Absolute DPHs and DPHs	All or none case series	Absolute number value or percentage analysis
2a	SR (with homogeneity) of cohort studies	SR (with homogeneity) of non-randomised cohort studies or individual level 2 studies	SR (with homogeneity) of Level 2 studies or SR (with homogeneity) of RCTs	SR (with homogeneity) of 2a and better studies	SR (with homogeneity) of Level 2C evidence studies
2b	Individual cohort study (with low risk of bias) (with confidence interval)	Non-randomised cohort study or follow-up study with low risk of bias, CDR, or published in a single publication	Individual study with low risk of bias, CDR, or published in a single publication	Retrospective cohort study or case-control study with low risk of bias, CDR, or published in a single publication	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2c	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2d	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2e	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2f	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2g	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2h	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2i	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2j	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2k	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2l	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2m	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2n	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2o	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2p	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2q	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2r	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2s	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2t	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2u	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2v	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2w	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2x	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2y	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
2z	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
3	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis
4	Case-control study	Case-control study	Case-control study	Case-control study	Analysis based on absolute number cases or events and relative risk (OR) of the outcome, or single studies, and including 95% CI and sensitivity analysis

## SIGN

**COMMUNITY MANAGEMENT OF LOWER RESPIRATORY TRACT INFECTION IN ADULTS**  
KEY TO EVIDENCE STATEMENTS AND GRADES OF RECOMMENDATIONS

Levels of evidence

- 1++ High quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias
- 1+ Well-conducted meta-analyses, systematic reviews, or RCTs with a low risk of bias
- 1- Meta-analyses, systematic reviews, or RCTs with a high risk of bias
- 2++ High quality systematic reviews of case control or cohort studies  
High quality case control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal
- 2+ Well-conducted case control or cohort studies with a low risk of confounding or bias and a moderate probability that the relationship is causal
- 2- Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal
- 3 Non-analytic studies, e.g. case reports, case series
- 4 Expert opinion

## GRADE

**Tabelle 2 Bedeutung der vier Stufen von Evidenz.**

Qualitätsstufe	Aktuelle Definition
Hoch	Wir sind sehr sicher, dass der wahre Effekt nahe bei dem Effektschätzer liegt.
Moderat	Wir haben mäßig viel Vertrauen in den Effektschätzer: Der wahre Effekt ist wahrscheinlich nahe bei dem Effektschätzer, aber es besteht die Möglichkeit, dass er relevant verschieden ist.
Niedrig	Unser Vertrauen in den Effektschätzer ist begrenzt: Der wahre Effekt kann durchaus relevant verschieden vom Effektschätzer sein.
Sehr niedrig	Wir haben nur sehr wenig Vertrauen in den Effektschätzer: Der wahre Effekt ist wahrscheinlich relevant verschieden vom Effektschätzer.

AWMF [2]: "Die Klassifikationsschemata von Oxford und SIGN heben auf die Ebene der Einzelstudien ab. Bei dem GRADE Ansatz wird die verfügbare Evidenz aus der Endpunkt- bzw. Outcome-Perspektive betrachtet (Bewertung der Gesamtheit der Studien als „body of evidence“ für jeden relevanten Endpunkt).

In einer Leitlinie sollte ein einheitliches Schema verwendet werden. Welches Schema verwendet wird, sollte in der Leitliniengruppe entschieden und im Leitlinienreport dokumentiert werden. Bei der Auswahl ist zu berücksichtigen, ob in größerem Umfang bereits vorhandene Leitlinien als Evidenzquellen herangezogen wurden (s. [Systematische Evidenzbasierung](#): Einführung, Leitlinienadaptation) und welches Schema/welche Schemata in diesen verwendet wurde/n. Werden in den verwendeten Quellleitlinien unterschiedliche Schemata angewendet, sollte geprüft werden, ob auf der Basis der dargelegten Einzelstudienbewertung die Überführung in ein einheitliches Schema möglich ist.“

Ziel der Evidenztabelle: Einschätzung der gegenwärtigen Evidenzlage zum Thema als Informationsgrundlage zur Formulierung und Graduierung von Empfehlungen

Beispiel für eine Evidenztabelle: s. <http://www.awmf.org/leitlinien/awmf-regelwerk/II-entwicklung/awmf-regelwerk-03-leitlinienentwicklung/II-entwicklung-recherche-literatur/II-entwicklung-evidenztabellen.html> und [http://www.awmf.org/fileadmin/user\\_upload/Leitlinien/AWMF-Regelwerk/Anhaenge/Anhang\\_06\\_Musterevidenztabelle.doc](http://www.awmf.org/fileadmin/user_upload/Leitlinien/AWMF-Regelwerk/Anhaenge/Anhang_06_Musterevidenztabelle.doc)

Im AWMF-Regelwerk Leitlinien: Recherche, Auswahl und methodische Bewertung von Literatur und deren Aufbereitung - Teil 3 wird die Aufbereitung und Erstellung von Evidenztabelle exemplarisch dargestellt, modifiziert nach Guidelines International Network – Evidence Tables Working Group: Das englische Original ist verfügbar unter: <http://www.g-i-n.net/document-store/working-groups-documents/etwg-documents/template-evidence-summary-intervention-studies>. (Stand: 20.07.2011)

### Graduierung der Empfehlungen

Die Graduierung der Empfehlungen erfolgt entsprechend den Vorgaben der AWMF (<http://www.awmf.org/leitlinien/awmf-regelwerk/II-entwicklung/awmf-regelwerk-03-leitlinienentwicklung/II-entwicklung-graduierung-der-empfehlungen.html>): „Die Graduierung der Empfehlungen basiert auf der identifizierten Evidenz, der klinischen Expertise und den Patientenpräferenzen und schließt damit auch explizit subjektiv wertende Elemente ein. Bei der Diskussion und Vergabe der Empfehlungsgrade sollten neben der zugrunde liegenden Evidenz konkret die folgenden Kriterien berücksichtigt werden:

- Konsistenz der Studienergebnisse
- Klinische Relevanz der Endpunkte und Effektstärken
- Nutzen-Risiko-Verhältnis
- Ethische, rechtliche, ökonomische Erwägungen
- Patientenpräferenzen
- Anwendbarkeit auf die Patientenzielgruppe und das deutsche Gesundheitssystem
- Umsetzbarkeit im Alltag / in verschiedenen Versorgungsbereichen.“

Die Darlegung der Qualität der Evidenz (Evidenzstärke) drückt die Belastbarkeit der Studienergebnisse und damit das Ausmaß an Sicherheit / Unsicherheit des Wissens aus. Die Darlegung der Empfehlungsgrade ist Ausdruck des Ergebnisses der Abwägung erwünschter bzw. unerwünschter Konsequenzen alternativer Vorgehensweisen. Dies kann dazu führen, dass Evidenz- und Empfehlungsstärken in begründeten Fällen voneinander abweichen können.

Leitlinie Evidenzbewertung in der Literatur

© BVA und DOG, alle Rechte vorbehalten

Berücksichtigt wird zudem das DELBI-Instrument Domäne 3 [3; 4] (Methodologische Exaktheit der Leitlinien-Entwicklung) mit dem Kriterium 11 (Berücksichtigung von gesundheitlichem Nutzen, Risiken und Nebenwirkungen) und dem Kriterium 12 (Begründung der angegebenen Empfehlungsgrade).

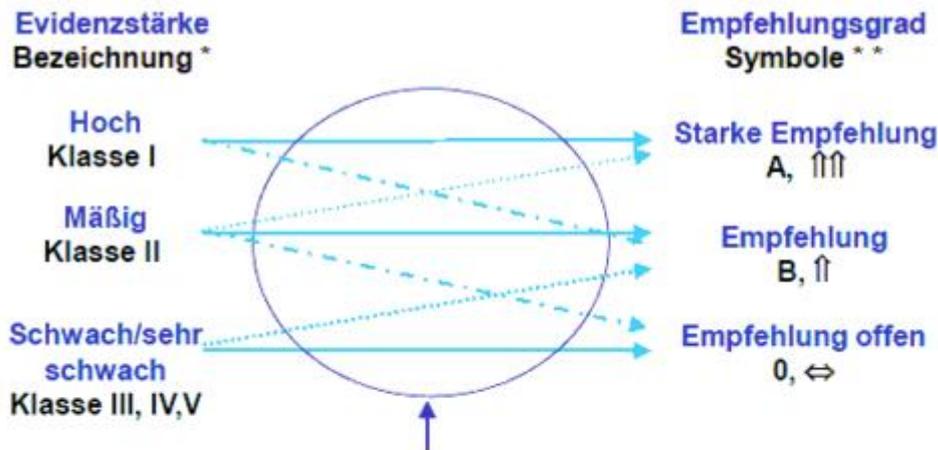
Schema 1: zur Graduierung von Empfehlungen [5]

<b>Empfehlungsgrad</b>	<b>Beschreibung</b>	<b>Syntax</b>
<b>A</b>	Starke Empfehlung	soll / soll nicht
<b>B</b>	Empfehlung	sollte / sollte nicht
<b>O</b>	Empfehlung offen	kann erwogen werden / kann verzichtet werden

Schema 2: Evidenzstärke – Empfehlungsgrad

„Die Evidenzbasierung einer Leitlinie setzt nicht voraus, dass jede in ihr enthaltene Einzelempfehlung mit einem hohen Evidenzlevel verknüpft ist. Auch systematisch und transparent erstellte evidenzbasierte Leitlinien enthalten in der Regel Empfehlungen, die auf einer schwachen Evidenzgrundlage beruhen.“ [6-7]

Von der Evidenz zur Empfehlung: Visualisierung der klinischen Beurteilung als Prozess der Kriteriengestützten Konsensusentscheidung



**Kriterien für die Graduierung (Konsensusaspekte):**

- Konsistenz der Studienergebnisse
- Klinische Relevanz der Endpunkte und Effektstärken
- Nutzen-Risiko-Verhältnis
- Ethische, rechtliche, ökonomische Erwägungen
- Patientenpräferenzen
- Anwendbarkeit, Umsetzbarkeit

- \* nach GRADE [BMJ 2004] [Bewertung der gesamten Evidenz (blau)]  
und Oxford Centre of Evidence based Medicine [Bewertung der Einzelstudien (Schwarz)]
- \*\* Empfehlungsgraduierung im Programm für Nationale Versorgungsleitlinien  
Die Empfehlungen werden nach Möglichkeit analog formuliert:  
Starke Empfehlung: „soll“; (abgeschwächte) Empfehlung: „sollte“;  
Negativ-Empfehlungen werden entweder rein sprachlich ausgedrückt („nicht“ / „kann verzichtet werden“) bei gleichen Symbolen oder sprachlich mit zusätzlich nach unten gerichteten Pfeilen;  
Offene Empfehlungen drücken eine Handlungsoption in Unsicherheit aus („kann erwogen werden“ / „kann verzichtet werden“).

1. ÄZQ, Methoden zur Aktualisierung von medizinischen Leitlinien,  
<http://www.aezq.de/mdb/edocs/pdf/schriftenreihe/schriftenreihe40.pdf>
2. AWMF, AWMF-Regelwerk Leitlinien: Recherche, Auswahl und methodische Bewertung von Literatur und deren Aufbereitung - Teil 2, <http://www.awmf.org/en/clinical-practice-guidelines/awmf-guidance/cpg-development/awmf-regelwerk-03-leitlinienentwicklung/ll-entwicklung-recherche-literatur.html>
3. [www.delbi.de](http://www.delbi.de)
4. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften  
Ärztliches Zentrum für Qualität in der Medizin  
Deutsches Instrument zur methodischen Leitlinien-Bewertung (DELBI),  
<http://www.leitlinien.de/mdb/edocs/pdf/literatur/delbi-fassung-2005-2006-domaene-8-2008.pdf>

5. GRADE working group, *Grading quality of evidence and strength of recommendations*. BMJ, 2004, [www.gradeworkinggroup.org](http://www.gradeworkinggroup.org)
6. IQWiG, *Systematische Leitlinienrecherche und -bewertung sowie Extraktion relevanter Empfehlungen für ein DMP chronischer Rückenschmerz*. <https://www.iqwig.de/de/projekte-ergebnisse/projekte/versorgungsqualitaet/v14-04-systematische-leitlinienrecherche-und-bewertung-sowie-extraktion-relevanter-empfehlungen-fur-ein-dmp-chronischer-rueckenschmerz.6255.html>
7. Ketola, E., M. Kaila, and M. Honkanen, *Guidelines in context of evidence*. Qual Saf Health Care, 2007. 16(4): p. 308-12.