

Prof. Dr. Dr. Holger Schünemann, M.Sc.
Berlin, 21. Oktober 2008

GRADE: Von der Evidenz zur Empfehlung

Beschreibung des Systems und Lösungsbeitrag zur Übertragbarkeit von Studienergebnissen

Education and debate

Grading quality of evidence and strength of recommendations

GRADE Working Group

Clinical guidelines are only as good as the evidence and judgments they are based on. The GRADE approach aims to make it easier for users to assess the judgments behind recommendations

RATING QUALITY OF EVIDENCE AND STRENGTH OF RECOMMENDATIONS

GRADE: an emerging consensus on rating quality of evidence and strength of recommendations

Guidelines are inconsistent in how they rate the quality of evidence and the strength of recommendations. This article explores the advantages of the GRADE system, which is increasingly being adopted by organisations worldwide

*Grade Working Group. CMAJ 2003, BMJ 2004, BMC 2004, BMC 2005, AJRCCM 2006, BMJ 2008

Übersicht

- Hintergrund und Prinzip „GRADE“
- Anwendung
- Übertragbarkeit von Studienergebnissen
 - „Directness“

Information durch Hierarchien von Evidenz und Empfehlungen

Systematische und explizite Ansätze können helfen:

- Information zu kommunizieren
 - Meinungsverschiedenheiten aufzulösen
 - gegen Irrtümer zu schützen
-
- Schlussfolgerungen zu ziehen über die:
 - Qualität der Evidenz
 - Stärke/Grad von Empfehlungen

Viele Systeme für ein Problem

Empfehlung für Vitamin K Antagonisten in Patienten mit Vorhofflimmern und Mitralklappenstenose

Evidenz	Empfehlung	Organisation
B	Class I	AHA
C+	1	ACCP
IV	C	SIGN

?



Über GRADE*

- Ziel: ein **einheitliches** System zur **transparenten** Beurteilung der Qualität der Evidenz und des Empfehlungsgrades zu entwickeln und die Bereiche zu untersuchen, bei denen das System Anwendung finden kann
- Arbeitsgruppe - seit 2000
- Wissenschaftler und Leitlinienentwickler mit Interesse die methodologischen Schwierigkeiten und Ungenauigkeiten aufzuarbeiten
- Evaluierung von vorhandenen Systemen

GRADE Working Group

David Atkins, chief medical officer^a

Dana Best, assistant professor^b

Martin Eccles, professor^d

Francoise Cluzeau, lecturer^x

Yngve Falck-Ytter, associate director^e

Signe Flottorp, researcher^f

Gordon H Guyatt, professor^g

Robin T Harbour, quality and information director^h

Margaret C Haugh, methodologistⁱ

David Henry, professor^j

Suzanne Hill, senior lecturer^j

Roman Jaeschke, clinical professor^k

Regina Kunx, Associate Professor

Gillian Leng, guidelines programme director^l

Alessandro Liberati, professor^m

Nicola Magrini, directorⁿ

James Mason, professor^d

Philippa Middleton, honorary research fellow^o

Jacek Mrukowicz, executive director^p

Dianne O'Connell, senior epidemiologist^q

Andrew D Oxman, director^f

Bob Phillips, associate fellow

Holger J Schünemann, professor^{g,5}

Tessa Tan-Torres Edejer, medical officer^t

David Tovey, Editor^y

Jane Thomas, Lecturer, UK

Helena Varonen, associate editor^u

Gunn E Vist, researcher^f

John W Williams Jr, professor^v

Stephanie Zaza, project director^w

a) Agency for Healthcare Research and Quality, USA

b) Children's National Medical Center, USA

c) Centers for Disease Control and Prevention, USA

d) University of Newcastle upon Tyne, UK

e) German Cochrane Centre, Germany

f) Norwegian Centre for Health Services, Norway

g) McMaster University, Canada

h) Scottish Intercollegiate Guidelines Network, UK

i) Fédération Nationale des Centres de Lutte Contre le Cancer, France

j) University of Newcastle, Australia

k) McMaster University, Canada

l) National Institute for Clinical Excellence, UK

m) Università di Modena e Reggio Emilia, Italy

n) Centro per la Valutazione della Efficacia della Assistenza Sanitaria, Italy

o) Australasian Cochrane Centre, Australia

p) Polish Institute for Evidence Based Medicine, Poland

q) The Cancer Council, Australia

r) Centre for Evidence-based Medicine, UK

s) National Cancer Institute, Italy

t) World Health Organisation, Switzerland

u) Finnish Medical Society Duodecim, Finland

v) Duke University Medical Center, USA

w) Centers for Disease Control and Prevention, USA

x) University of London, UK

Y) BMJ Clinical Evidence, UK

GRADE Nutzer/Anwender

- World Health Organization
- National Institute Clinical Excellence (NICE)
- Agency for Health Care Research and Quality (AHRQ)
- Canadian Agency for Drugs and Technology in Health (CADTH)
- Allergic Rhinitis in Asthma Guidelines (ARIA)
- American Thoracic Society
- American College of Chest Physicians
- UpToDate
- British Medical Journal
- American College of Physicians
- Cochrane Collaboration
- European Society of Thoracic Surgeons
- Clinical Evidence
- Many other organizations

Was wird beurteilt?

1. Gesamtschau/Qualität der Evidenz für eine Fragestellung
 - endpunktspezifisch
 - Wahrscheinlichkeit das systematische Fehler vorliegen oder Übertragbarkeit eingeschränkt ist
 - Kein Instrument für die Beurteilung einzelner Studien
2. Stärke/Grad der Empfehlung
 - stark und schwach/bedingt
 - Qualität nur ein Faktor

Prozess der Leitlinienentwicklung

Einrichtung einer LL-Gruppe

Schlüsselfragen und systematische Reviews

Evidenztabellen/-profile

- Qualität der Evidenz für jeden Endpunkt
(inkl. Gesamtqualität der Evidenz)
- Relative Bedeutung der einzelnen Endpunkte
- Abwägung zwischen Nutzen und Schaden // (\pm Kosten)
 - Stärke der Empfehlung

Implementierung und Evaluation

Prozess der Leitlinienentwicklung

Einrichtung einer LL-Gruppe

Schlüsselfragen und systematische Reviews

Evidenztabellen/-profile



Summary
of Findings

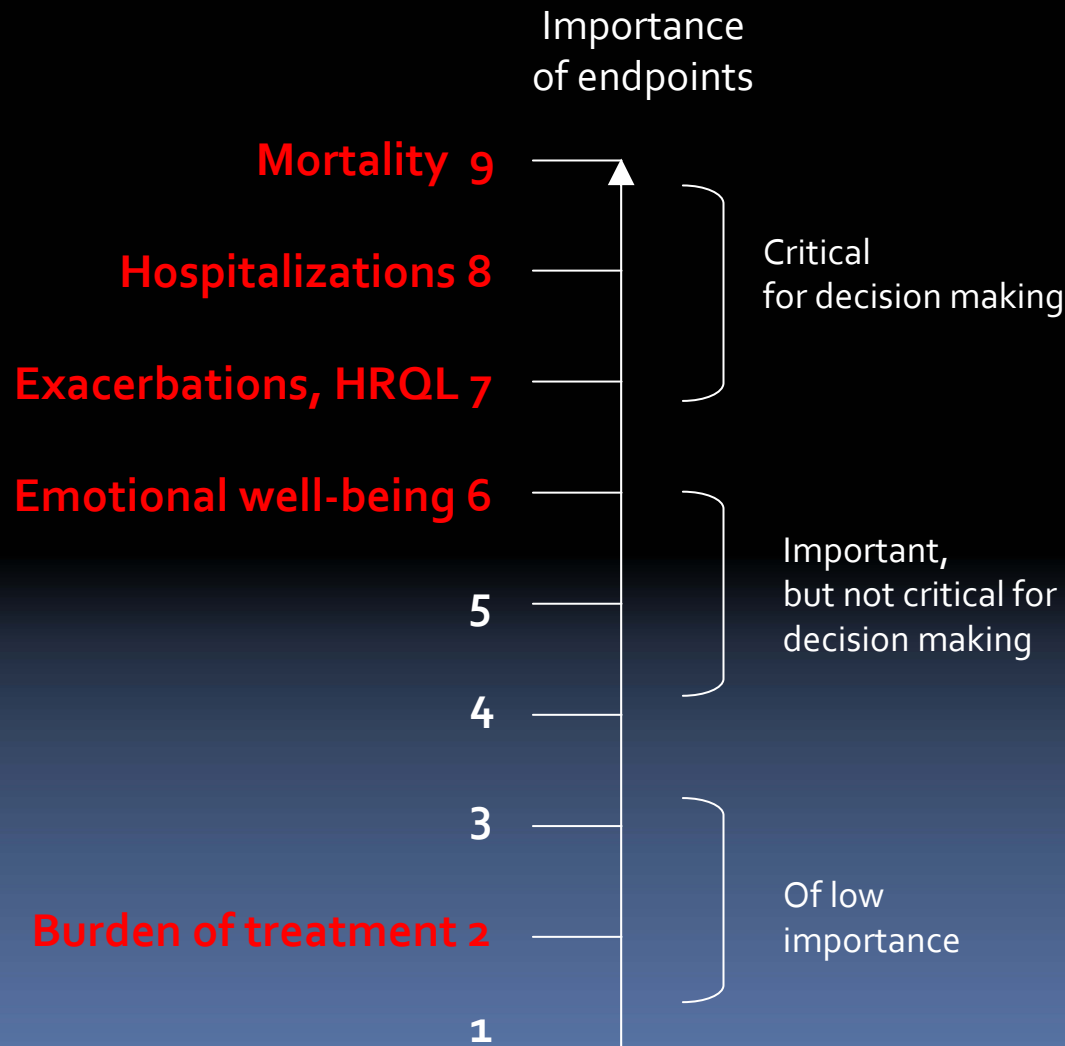
- Qualität der Evidenz für jeden Endpunkt (inkl. Gesamtqualität der Evidenz)
- Relative Bedeutung der einzelnen Endpunkte
- Abwägung zwischen Nutzen und Schaden // (\pm Kosten)
 - Stärke der Empfehlung

Implementierung und Evaluation

Klinische Fragestellung

Population:	Patienten mit COPD
Intervention:	kombinierte aerosole (ICS+LABA)
Comparison:	keine derartige Therapie
Outcomes:	↓ Mortalität und Exazerbationen ↑ Nebenwirkungen/Pneumonien Verbesserung HRQL?

Hierarchie von Endpunkten in der COPD



Evidenzprofile

GRADE Evidence Profile

Author(s): Santesso, Schünemann, Nannini, Cates, Lasserson, Poole

Date: 2007-08-06

Question: Should corticosteroid and long-acting beta-agonist in one inhaler vs no treatment be used for moderate and severe chronic obstructive pulmonary disease?

Bibliography: Nannini LJ, Cates CJ, Lasserson TJ, Poole P. Cochrane Database of Systematic Reviews 2007, Issue 4.

Quality assessment							Summary of findings					Importance for decision making
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	No of patients		Effect		Quality	
							corticosteroid and long-acting beta-agonist in one inhaler ³	no treatment	Relative (95% CI)	Absolute		
Exacerbation rate (follow-up 3 years)												
5	randomised trial	no serious limitations	no serious inconsistency	serious ¹	no serious imprecision	none	2191	2035	Rate Ratio 0.74 (0.69 to 0.79)	1 less exacerbation per 3 years per patient	⊕⊕⊕○ MODERATE	critical
Hospitalisations - not reported												
0	-	-	-	-	-	none	0/0	0/0	-	-		
Mortality (follow-up 3 years)												
7	randomised trial	no serious limitations	no serious inconsistency	no serious indirectness	serious ²	none	209/2946	255/2806 (9.1%)	RR 0.80 (0.65 to 0.96)	18 fewer per 1000	⊕⊕⊕○ MODERATE	critical
Quality of Life (follow-up 3 years; measured with: St. George's Respiratory Questionnaire; range of scores: 0-100; Better indicated by lower scores)												
4	randomised trial	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	1788	1558	-	MD -2.90 (-3.61 to -2.18)	⊕⊕⊕⊕ HIGH	critical
Pneumonia (follow-up 3 years)												
8	randomised trial	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	333/2927	196/2812 (6.9%)	RR 1.80 (1.51 to 2.21)	55 more per 1000	⊕⊕⊕⊕ HIGH	critical
Any adverse events (follow-up 3 years)												
8	randomised trial	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	2215/2808	2116/2685	RR 1.01 (0.96 to 1.27)	0 more per 1000	⊕⊕⊕⊕ HIGH	critical

¹ Withdrawal of participants with severe frequent exacerbations may limit inference for severe patients.

² Sparse data.

³ Both long-acting beta-agonists and inhaled corticosteroids can be used in combination for the treatment of chronic obstructive pulmonary disease. Of the 11 included studies, two evaluated fluticasone/salmeterol at 250 mcg/50 mcg twice daily and seven at 500 mcg/50 mcg twice daily; and two evaluated budesonide/formoterol at 320 mcg/9 mcg twice daily.

Quality assessment						
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations
Exacerbation rate (follow-up 3 years)						
5	randomised trial	no serious limitations	no serious inconsistency	serious ¹	no serious imprecision	none
Hospitalisations - not reported						
0	-	-	-	-	-	none
Mortality (follow-up 3 years)						
7	randomised trial	no serious limitations	no serious inconsistency	no serious indirectness	serious ²	none
Quality of Life (follow-up 3 years; measured with: St. George's Respiratory Questionnaire; range of score 0-100)						
4	randomised trial	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none
Pneumonia (follow-up 3 years)						
8	randomised trial	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none
Any adverse events (follow-up 3 years)						
8	randomised trial	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none

¹ Withdrawal of participants with severe frequent exacerbations may limit inference for severe patient

² Sparse data.

³ Both long-acting beta-agonists and inhaled corticosteroids can be used in combination for the treatment of asthma. The study compared fluticasone/salmeterol at 250 mcg/50 mcg twice daily and seven at 500 mcg/50 mcg twice daily; and

Evidenzprofile

GRADE Evidence Profile

Author(s): Santesso, Schünemann, Nannini, Cates, Lasserson, Poole

Date: 2007-08-06

Question: Should corticosteroid and long-acting beta-agonist in one inhaler vs no treatment be used for moderate and severe chronic obstructive pulmonary disease?

Bibliography: Nannini LJ, Cates CJ, Lasserson TJ, Poole P. Cochrane Database of Systematic Reviews 2007, Issue 4.

Quality assessment							Summary of findings					Importance for decision making
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	No of patients		Effect		Quality	
							corticosteroid and long-acting beta-agonist in one inhaler ³	no treatment	Relative (95% CI)	Absolute		
Exacerbation rate (follow-up 3 years)												
5	randomised trial	no serious limitations	no serious inconsistency	serious ¹	no serious imprecision	none	2191	2035	Rate Ratio 0.74 (0.69 to 0.79)	1 less exacerbation per 3 years per patient	⊕⊕⊕○ MODERATE	critical
Hospitalisations - not reported												
0	-	-	-	-	-	none	0/0	0/0	-	-		
Mortality (follow-up 3 years)												
7	randomised trial	no serious limitations	no serious inconsistency	no serious indirectness	serious ²	none	209/2946	255/2806 (9.1%)	RR 0.80 (0.65 to 0.96)	18 fewer per 1000	⊕⊕⊕○ MODERATE	critical
Quality of Life (follow-up 3 years; measured with: St. George's Respiratory Questionnaire; range of scores: 0-100; Better indicated by lower scores)												
4	randomised trial	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	1788	1558	-	MD -2.90 (-3.61 to -2.18)	⊕⊕⊕⊕ HIGH	critical
Pneumonia (follow-up 3 years)												
8	randomised trial	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	333/2927	196/2812 (6.9%)	RR 1.80 (1.51 to 2.21)	55 more per 1000	⊕⊕⊕⊕ HIGH	critical
Any adverse events (follow-up 3 years)												
8	randomised trial	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	2215/2808	2116/2685	RR 1.01 (0.96 to 1.27)	0 more per 1000	⊕⊕⊕⊕ HIGH	critical

¹ Withdrawal of participants with severe frequent exacerbations may limit inference for severe patients.

² Sparse data.

³ Both long-acting beta-agonists and inhaled corticosteroids can be used in combination for the treatment of chronic obstructive pulmonary disease. Of the 11 included studies, two evaluated fluticasone/salmeterol at 250 mcg/50 mcg twice daily and seven at 500 mcg/50 mcg twice daily; and two evaluated budesonide/formoterol at 320 mcg/9 mcg twice daily.

Summary of findings					
No of patients		Effect		Quality	Importance for decision making
corticosteroid and long-acting beta-agonist in one inhaler ³	no treatment	Relative (95% CI)	Absolute		
2191	2035	Rate Ratio 0.74 (0.69 to 0.79)	1 less exacerbation per 3 years per patient	⊕⊕⊕O MODERATE	critical
0/0	0/0	-	-		
209/2946	255/2806 (9.1%)	RR 0.80 (0.65 to 0.96)	18 fewer per 1000	⊕⊕⊕O MODERATE	critical
Notes: 0-100; Better indicated by lower scores)					
1788	1558	-	MD -2.90 (-3.61 to -2.18)	⊕⊕⊕⊕ HIGH	critical
333/2927	196/2812 (6.9%)	RR 1.80 (1.51 to 2.21)	55 more per 1000	⊕⊕⊕⊕ HIGH	critical
2215/2808	2116/2685	RR 1.01 (0.96 to 1.27)	0 more per 1000	⊕⊕⊕⊕ HIGH	critical

GRADE - Qualität der Evidenz

„Ein Gradmesser für die Zuversicht, dass ein ermittelter Effekt korrekt ist.“

Qualität

- Hoch (Randomisierte Studien)
- Mittel
- Niedrig (Beobachtungsstudien)
- Sehr niedrig

Experten Meinung = ~~Evidenz~~

GRADE - Qualität der Evidenz

Herabstufung durch:

- Bias-anfällige(s) Studiendesign oder Durchführung (RoB)
 - „Blinding, Randomisierung, Concealment, Follow-up, ITT“, Studienunterbrechung wg. pos. Effekt, selective outcome reporting
- Heterogenität der Resultate/Evidenz (Inconsistency)
- Unpräzise Datenlage (imprecise data)
- „Publication Bias“ (Publikation)
- Geringe Vergleichbarkeit/Übertragbarkeit der Evidenz (indirekte Evidenz)
 - Vergleich, Population, Intervention, Endpunkte

Geringe Vergleichbarkeit/ Übertragbarkeit der Evidenz

- 1) Indirekter Vergleich
- 2) Population, Intervention, Outcome

Geringe Vergleichbarkeit/ Übertragbarkeit der Evidenz

1) Indirekter Vergleich

Thromboseprophylaxe im Krankenhaus

2 x versus 3 x tägliche Gabe von Heparin

Viele Studien, die 2 x oder 3 x tägliche Gabe mit Placebo verglichen haben, aber kein direkter Vergleich

Verringert unser Vertrauen in den direkten Vergleich

Geringe Vergleichbarkeit/ Übertragbarkeit der Evidenz

2) Population, Intervention, Outcome

Fragestellung

Oseltamivir zur Behandlung der
Vogelgrippe durch Influenza A(H5N1)
Virus

Indirekte Evidenz

Population: Randomisierte Studien mit
Oseltamivir sind vorhanden, aber für die
gewöhnliche Grippe, nicht für die Vogelgrippe

Sigmoidoskopie "screening" zur
Prävention von Maglinomen des
Kolons

Intervention: Randomisierte Studien "fecal occult
blood screening" indirekt für Sigmoidoskopie

Heparin zur Thromboseprophylaxe

Indirekte "outcomes": asymptomatische versus
symptomatische tiefe Venenthrombose

Surrogate and patient-important outcomes for phosphate lowering drugs in patients with renal failure and hyperphosphatemia

Surrogate outcomes

Indirect

(Lower the quality of evidence by one level)

Very indirect

(Lower the quality of evidence by two levels)

Patient-important outcomes

Myocardial infarction

Coronary calcification

Fractures

Bone density

Pain due to soft tissue calcification

Soft tissue calcification

Measures of calcium/phosphate metabolism

GRADE - Qualität der Evidenz

Heraufstufung durch:

- Vorhandensein einer starken Assoziation
 - $RR > 2$ ($< 0,5$), wenn mehr als zwei Studien ohne plausible Confounder und mit konsistenten Ergebnissen vorliegen
 - $RR > 5$ ($< 0,2$)
- Vorhandensein einer Dosis-Wirkungs-Beziehung
- Alle verbleibenden, plausiblen „Confounder“ haben den beobachteten Effekt bereits reduziert oder einen abwesenden Effekt möglicherweise verstärkt

GRADE - Kriterien zur Bewertung der Qualität

Qualität der Evidenz	Studiendesign	Herabstufen falls*	Hinaufstufen falls *
Hoch	Randomisierte Studie	Studienqualität: Schwerwiegende Einschränkungen -1 Sehr schwerwiegende Einschränkungen -2	die Assoziation stark, ohne plausible Confounder, konsistente u. direkte** Evidenz +1 sehr stark, ohne Einschränkung d. Validität, konsistente und direkte** Evidenz +2
Mittel		Widersprüchliche Effekte Wesentl. inkonsistente Effekte -1 oder -2 Direktheit** Einige Unsicherheit -1 Große Unsicherheit -2	Dosis-Wirkungsbeziehung Evidenz für eine Dosis-Wirkungsbeziehung +1
Niedrig	Beobachtungs-Studie	Vorhandene Daten Wenige oder ungenaue Daten -1 oder -2	Confounder Alle plausiblen Confounder hätten den beobachteten Effekt verringert oder einen abwesenden Effekt verstaerkt +1
Sehr niedrig		Reporting Bias Hohe Wahrscheinlichkeit für Reporting Bias -1 oder -2	

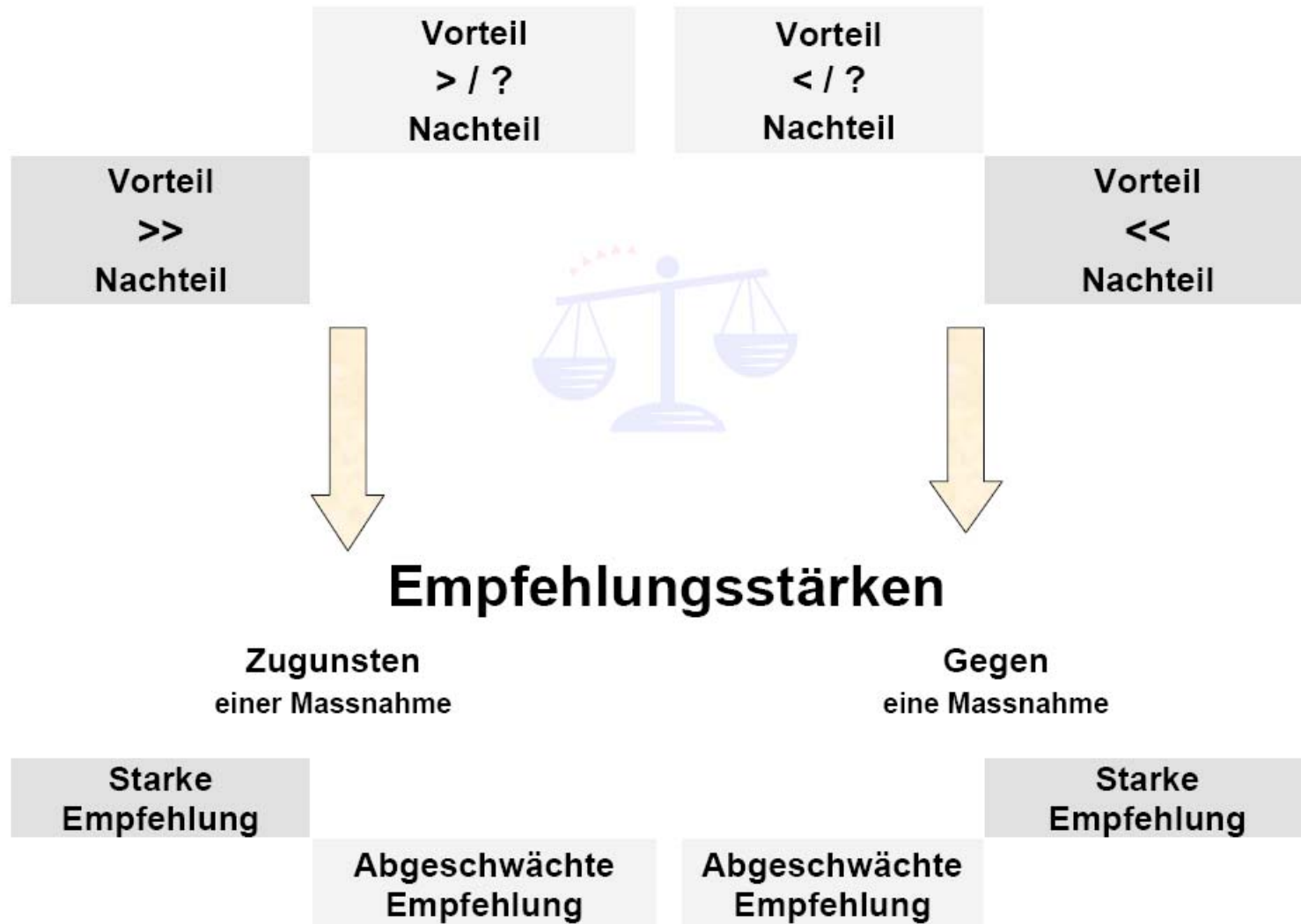
GRADE - Interpretationen

Hoch ⊕ ⊕ ⊕ ⊕	Es ist unwahrscheinlich, dass weitere Forschung unser Vertrauen in den beobachteten Effekt ändern wird
Mittel ⊕ ⊕ ⊕ ○	Weitere Forschung hat <u>vermutlich</u> einen wichtigen Einfluss auf unser Vertrauen in den Effekt und <u>kann</u> zu einer <u>Änderung</u> führen
Niedrig ⊕ ⊕ ○ ○	Es ist <u>sehr wahrscheinlich</u> , dass weitere Forschung unser Vertrauen in den beobachteten Effekt <u>stark beeinflussen</u> mit <u>Änderung</u> des bisher beobachteten Effekts
Sehr niedrig ⊕ ○ ○ ○	Jeder beobachtete Effekt ist sehr unsicher

GRADE - Empfehlungsstärke

„Die Stärke der Empfehlung ist ein Gradmesser für die Zuversicht, dass das Umsetzen der Empfehlung mehr erwünschte als unerwünschte Effekte hat“

Abwägen der Vor- und Nachteile



Entscheidung über den Stärkegrad einer Empfehlung

Faktoren die den Stärkegrad einer Empfehlung beeinflussen und zu einer schwachen Empfehlung führen können	Erklärung
Qualität der Evidenz [Niedrigere Qualität der Evidenz]	Unsicherheit bezüglich der Grösse des (relativen) Effekts (Vor- und Nachteile)
Ausmass des Nutzens/Schadens [Unsicherheit bezüglich des Abwägens der Vor- und Nachteile]	Unsicherheit bezüglich des Grundrisikos, Prävalenz einer Erkrankung, die den (absoluten) Effekt beeinflussen könnten.
Werte und Präferenzen [Unsicherheit bezüglich der relativen Werte, die den outcomes beigemessen werden]	Unsicherheit bezüglich der relativen Werte der Vor- und Nachteile (Nutzen und Schaden) für die betroffenen Patienten oder die Bevölkerung oder grosse Unterschiede, wie Individuen die outcomes bewerten. Diese Unsicherheit und Unterschiede können das Abwägen beeinflussen
Kosten (resources) [Unsicherheit ob die Vorteile kosteneffizient sind]	Unsicherheit oder Abwesenheit von Informationen bezüglich der Kosten oder ob der Ressourcenaufwand durch den angenommenen Nutzen ausgeglichen wird.

GRADE - Die Empfehlung

<p>Stark Do it/Don't do it We recommend (not)</p>	<ul style="list-style-type: none">• Die <u>meisten</u> „informierten“ Personen teilen diese Einschätzung• „Health Policies“• „Qualitätsmarker“
<p>Schwach/bedingt Probably do it/ do not do it We suggest (not)</p>	<ul style="list-style-type: none">• Die <u>Mehrheit</u> „informierter“ Personen teilt diese Einschätzung, eine <u>wesentliche Minderheit</u> hätte eine andere Einschätzung• Genauere Erörterung• „Decision aids“

Zusammenfassung

Herausforderungen beim Zuweisen der Gradmesser

Beurteilungen sind immer nötig

Untertrennung von Handlungsempfehlung und
Qualitätsbeurteilung

Abwägen von Nutzen und Schaden/Kosten

Grösse der Effekte, Präzision, Werte

Studiendesign, Studiendurchführung, Konsistenz,
Übertragbarkeit (PICO), Publikationsbias

GRADE Working Group aktiv in Verbreitung und Dialog