Holger Schünemann, MD, PhD
Professor of Clinical Epidemiology & Biostatistics, McMaster University
Associate Professor of Medicine, SUNY, Buffalo

Porto Alegre, September 24, 2008
World Congress of Epidemiology

GRADE

GUIDELINE DEVELOPMENT METHODOLOGY USING THE GRADE APPROACH

Content

- Background and rationale for revisiting guideline methodology and GRADE
- GRADE approach
  - Quality of evidence
  - Strength of recommendations
- Examples of Guidelines based on the GRADE approach

Content

- Background and rationale for revisiting guideline methodology and GRADE
- GRADE approach
  - Quality of evidence
  - Strength of recommendations
- Examples of Guidelines based on the GRADE approach

Guidelines

Systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances

Institute of Medicine, 1992

Confidence in evidence

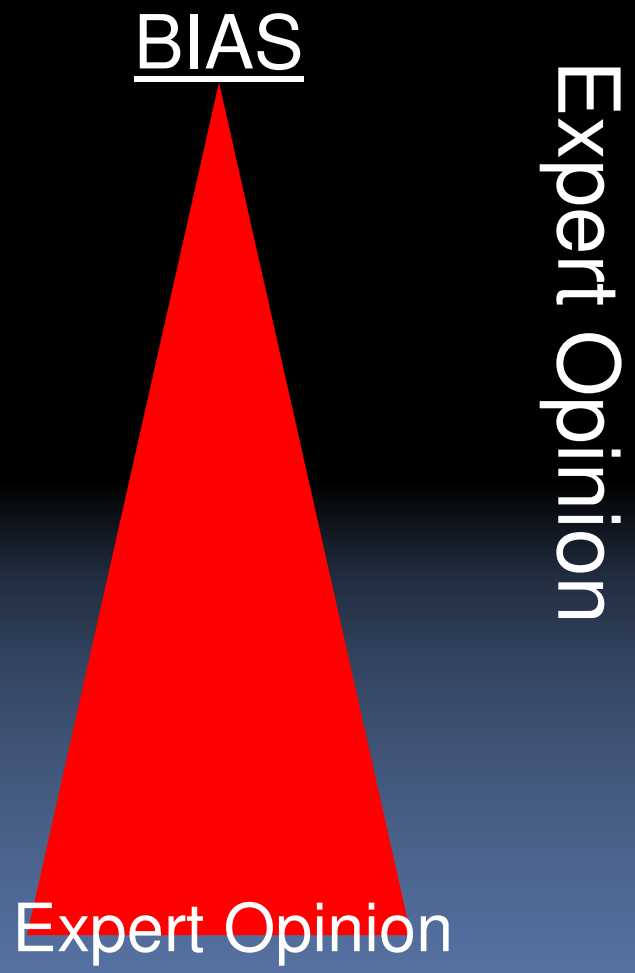
- There always is evidence
  - “When there is a question there is evidence”
- Research evidence alone is never sufficient to make a clinical decision
- Better research ⇒ greater confidence in the evidence and decisions

Hierarchy of evidence

STUDY DESIGN

- Randomized Controlled Trials
- Cohort Studies and Case Control Studies
- Case Reports and Case Series, Non-systematic observations

Expert Opinion



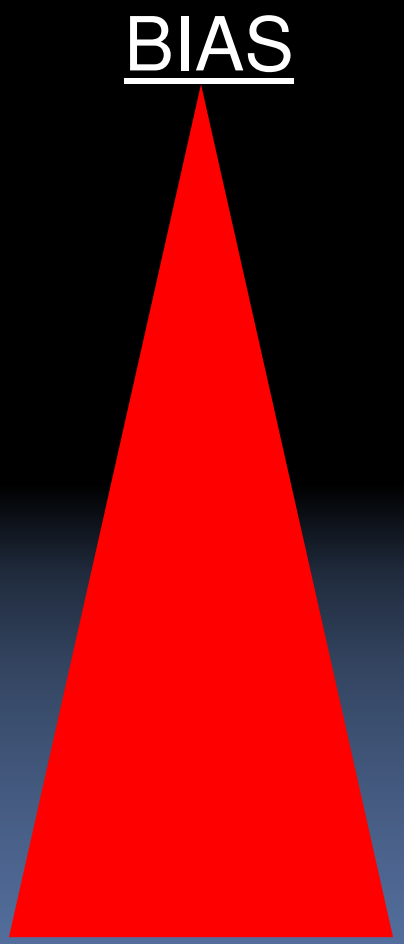
What about the following?

- Concealment of randomization
- Blinding (who is blinded in a double blinded trial?)
- Intention to treat analysis and its correct application
- Why trials stopped early for benefit overestimate treatment effects?
- P-values and confidence intervals

Hierarchy of evidence

STUDY DESIGN

- Randomized Controlled Trials
- Cohort Studies and Case Control Studies
- Case Reports and Case Series, Non-systematic observations



Expert Opinion

Reasons for grading evidence?

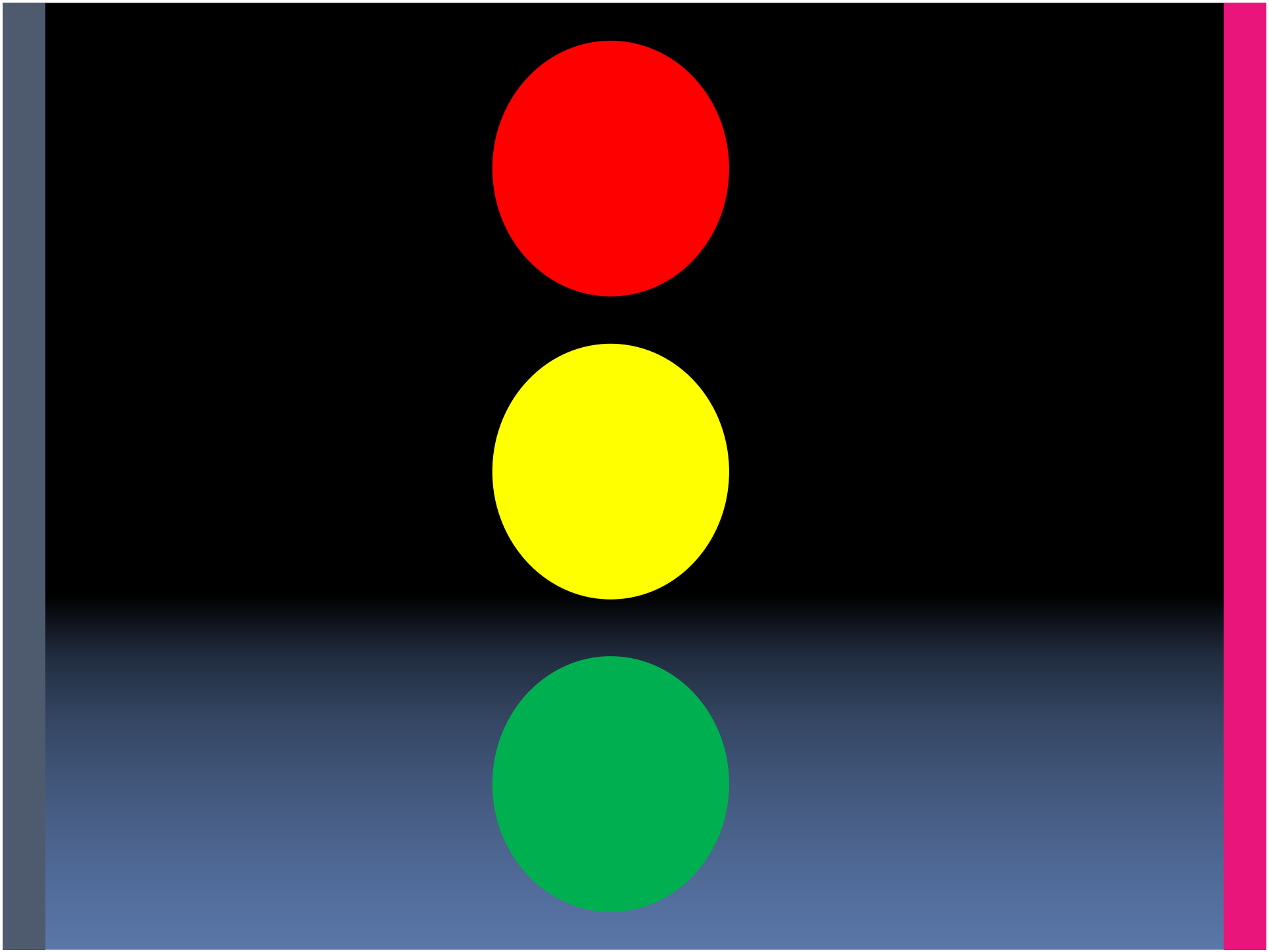
- People draw conclusions about the
  - quality of evidence and strength of recommendations
- Systematic and explicit approaches can help
  - protect against errors, resolve disagreements
  - communicate information and fulfil needs
- Change practitioner behavior
- However, wide variation in approaches

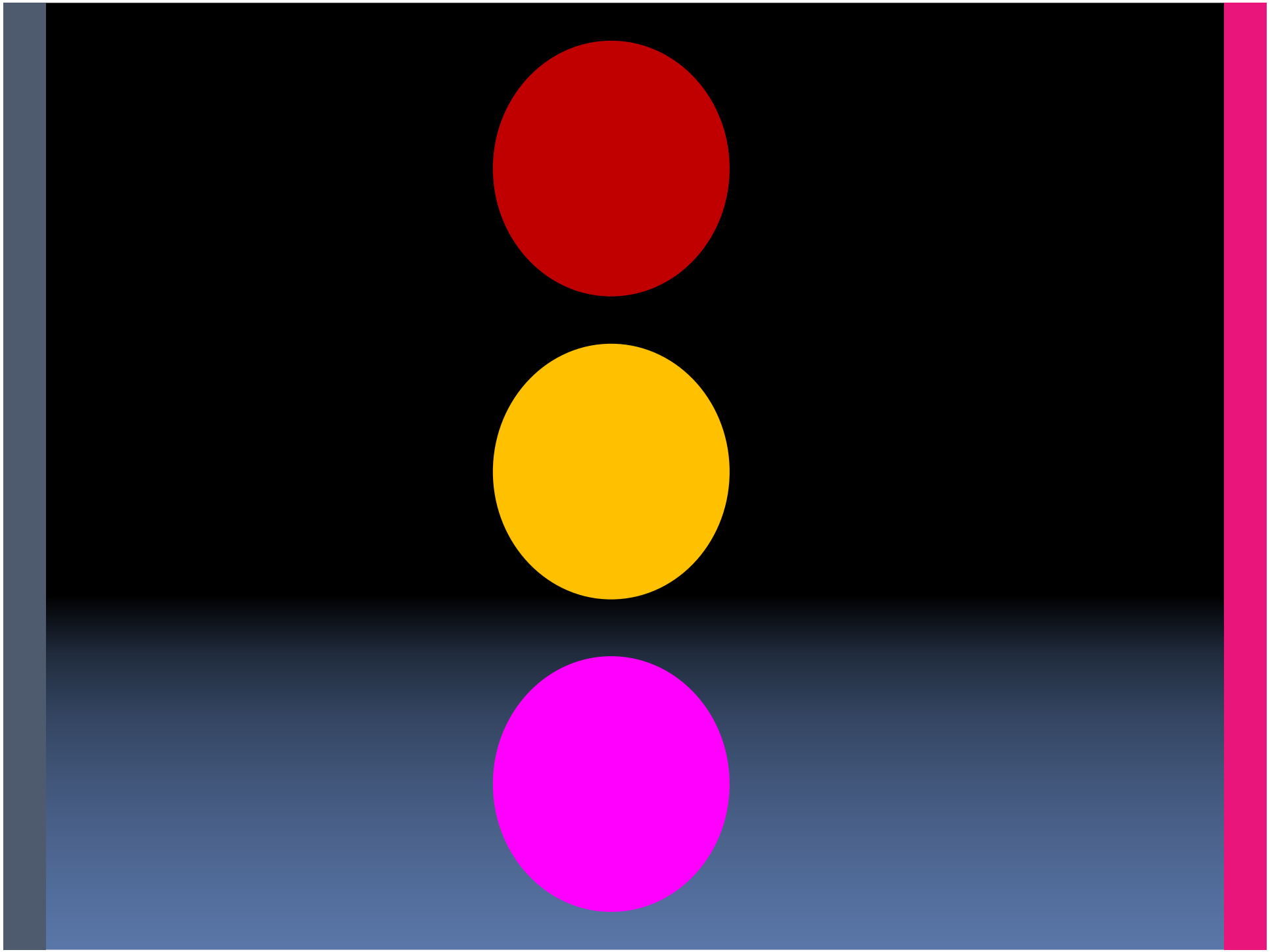
GRADE working group. BMJ. 2004 & 2008

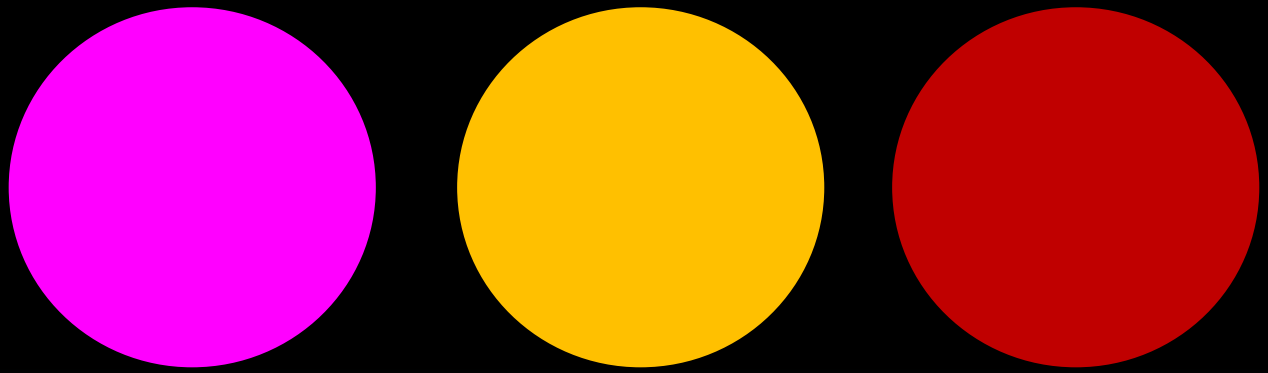
Which grading system?

Recommendation for use of oral anticoagulation in patients with atrial fibrillation and rheumatic mitral valve disease

Evidence	Recommendation	Organization
B	Class I	AHA
A	1	ACCP
IV	C	SIGN







Content

- Background and rationale for revisiting guideline methodology and GRADE
- GRADE approach
  - Quality of evidence
  - Strength of recommendations
- Examples of Guidelines based on the GRADE approach

Grades of **R**ecommendation **A**ssessment,
Development and **E**valuation

GRADE

WORKING GROUP

RATING QUALITY OF EVIDENCE AND STRENGTH OF RECOMMENDATIONS

GRADE

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

BMJ | 26 APRIL 2008 | VOLUME 336

CMAJ 2003, BMJ 2004, BMC 2004, BMC 2005, AJRCCM 2006, Chest 2006, BMJ 2008

About GRADE

- Since 2000
- Researchers/guideline developers with interest in methodology
- Aim: to develop a common, transparent and sensible system for grading the quality of evidence and the strength of recommendations
- Evaluation of existing systems

GRADE Uptake

- World Health Organization
- UpToDate
- British Medical Journal
- American College of Physicians
- Cochrane Collaboration
- National Institute of Health and Clinical Excellence, UK (NICE)
- American Thoracic Society
- European Society of Thoracic Surgeons
- Clinical Evidence
- American College of Chest Physicians
- Agency for Health Care Research and Quality (AHRQ)
- Allergic Rhinitis in Asthma Guidelines (ARIA)
- Over 20 organizations

It begins with the question: case scenario

A 13 year old girl who lives in rural Indonesia presented with flu symptoms and developed severe respiratory distress over the course of the last 2 days. She required intubation. The history reveals that she shares her living quarters with her parents and her three siblings. At night the family's chicken stock shares this room too and several chicken had died unexpectedly a few days before the girl fell sick.

Relevant clinical question?

Example from a *not so common* disease

Clinical question:

Population: Avian Flu/influenza A (H5N1) patients

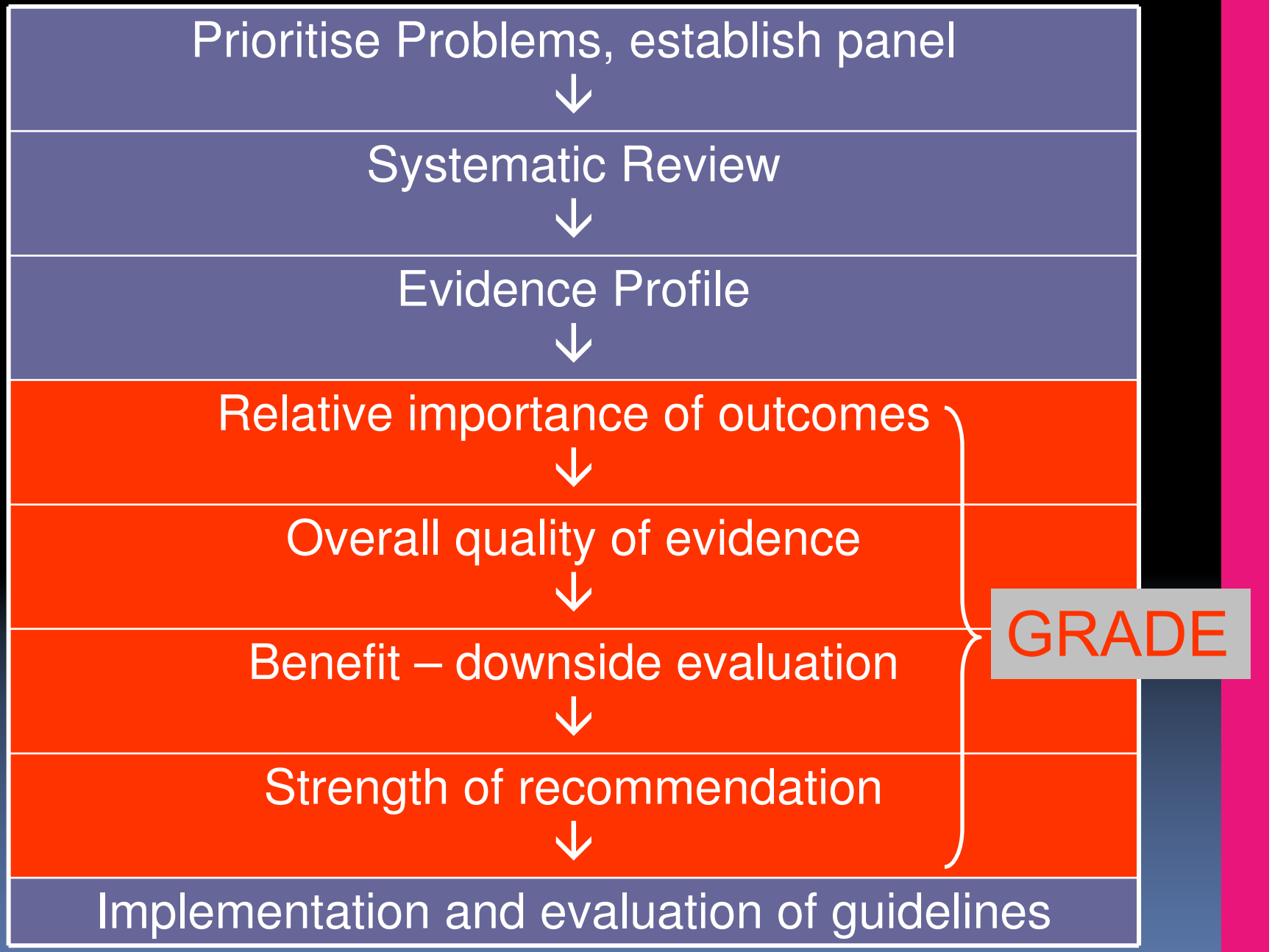
Intervention: Oseltamivir

Comparison: No pharmacological intervention

Outcomes: Mortality, hospitalizations, resource use, adverse outcomes, antimicrobial resistance

Schunemann et al. The Lancet ID, 2007

process


Prioritise Problems, establish panel
Systematic Review
Evidence Profile
Relative importance of outcomes
Overall quality of evidence
Benefit – downside evaluation
Strength of recommendation
Implementation and evaluation of guidelines
GRADE

The GRADE approach

Clear separation of 2 issues:

1) 4 categories of quality of evidence: very low, low, moderate, or high quality?
   - methodological quality of evidence
   - likelihood of bias
   - by outcome

2) Recommendation: 2 grades - weak or strong (for or against)?
   - Quality of evidence only one factor

*www.GradeWorkingGroup.org

GRADE Quality of Evidence

"Extent to which one is confident that the estimate of effect is adequate to support a decision"

- high: considerable confidence in estimate of effect.
- moderate: further research likely to have impact on confidence in estimate, may change estimate.
- low: further research is very likely to impact on confidence, likely to change the estimate.
- very low: any estimate of effect is very uncertain

Determinants of quality

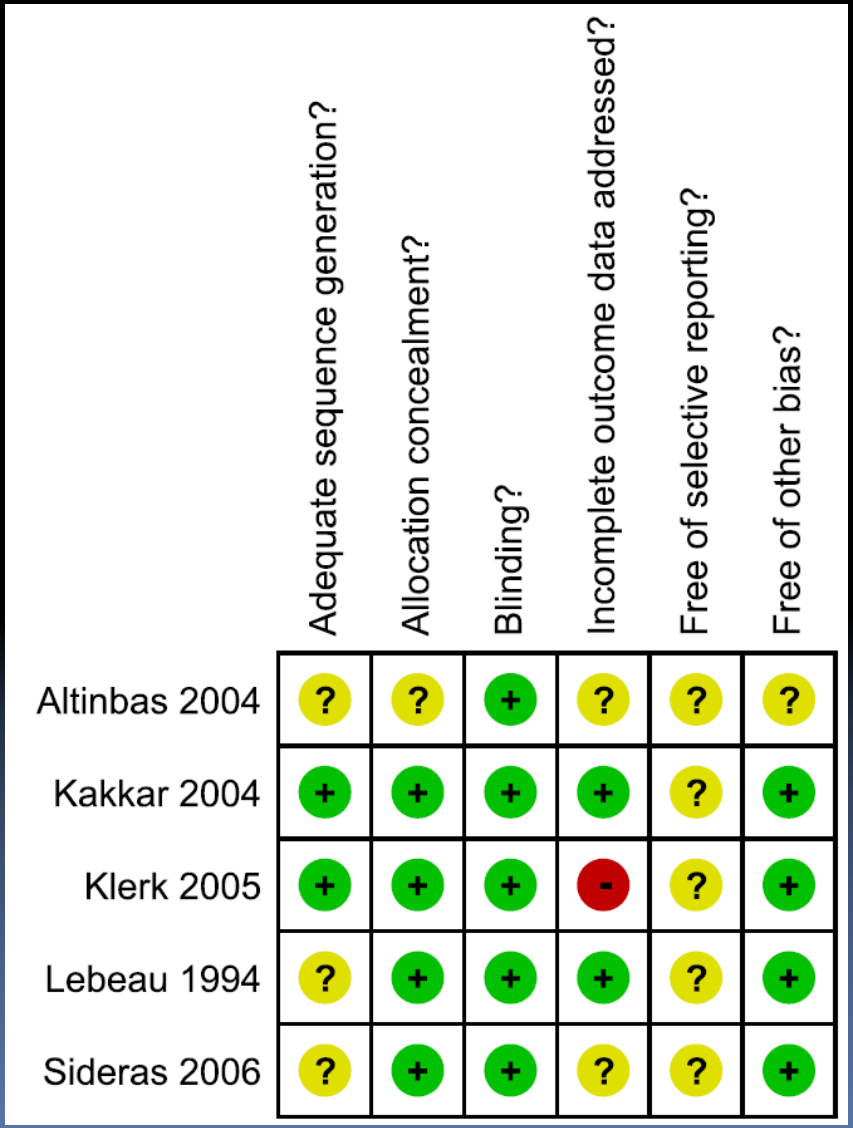
- RCTs start high
- observational studies start low
- 5 factors that can lower quality
  1. limitations of detailed design and execution
  2. inconsistency
  3. indirectness
  4. reporting bias
  5. Imprecision
- 3 factors can increase quality
  1. large magnitude of effect
  2. all plausible confounding may be working to reduce the demonstrated effect or increase the effect if no effect was observed
  3. dose-response gradient

1. Design and Execution

- limitations
  - Randomization
  - lack of concealment
  - intention to treat principle violated
  - inadequate blinding
  - loss to follow-up
  - early stopping for benefit
  - selective reporting of outcomes

Heparin or vitamin K antagonists for survival in patients with cancer

Detailed study design and execution: major bleeding



2. Inconsistency of results

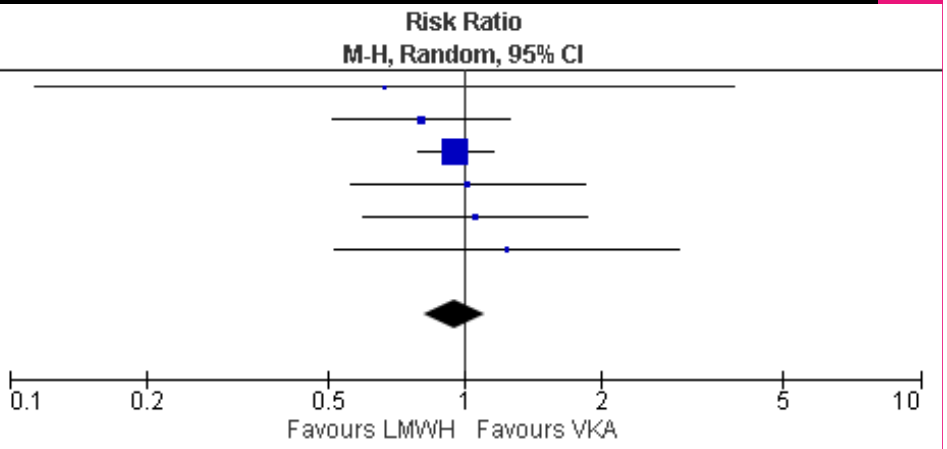
- Look for explanation for inconsistency
  - patients, intervention, comparator, outcome, methods
- Judgment
  - variation in size of effect
  - overlap in confidence intervals
  - statistical significance of heterogeneity
  - I^2

Heparin or vitamin K antagonists for survival in patients with cancer

Study or Subgroup	LMWH Events	LMWH Total	VKA Events	VKA Total	Weight	Risk Ratio M-H, Random, 95% CI
Cesarone 2003	2	96	3	96	0.8%	0.67 [0.11, 3.90]
Meyer 2002	22	71	29	75	12.0%	0.80 [0.51, 1.26]
Lee 2003	130	336	136	336	69.5%	0.96 [0.79, 1.15]
Deitcher 2006	22	67	11	34	6.9%	1.01 [0.56, 1.84]
Hull 2006	20	100	19	100	7.6%	1.05 [0.60, 1.85]
Lopez Beret 2001	7	17	6	18	3.2%	1.24 [0.52, 2.94]
Total (95% CI)		**687**		**659**	**100.0%**	**0.95 [0.81, 1.11]**
Total events	203		204			

Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 1.24$, df = 5 (P = 0.94); $I^2 = 0\%$

Test for overall effect: Z = 0.62 (P = 0.53)



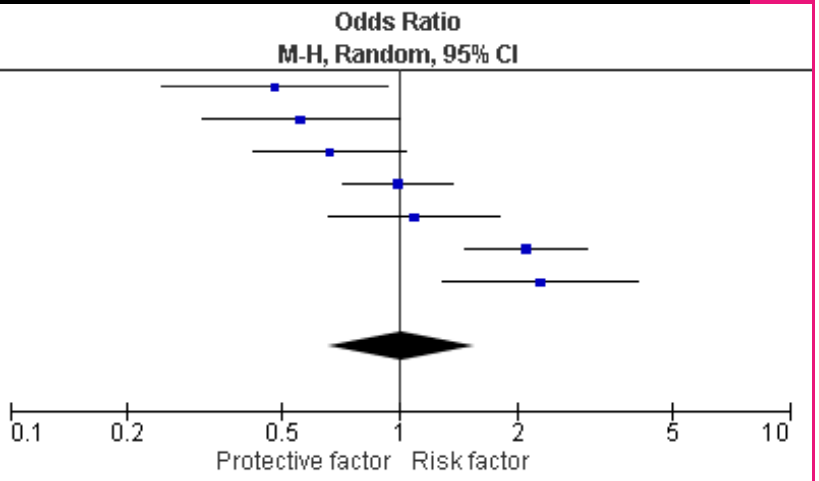
Akl E, Barba M, Rohilla S, Terrenato I, Sperati F, Schünemann HJ. "*Anticoagulation for the long term treatment of venous thromboembolism in patients with cancer*". Cochrane Database Syst Rev. 2008 Apr 16;(2):CD006650.

Non-steroidal drug use and risk of pancreatic cancer

Study or Subgroup	ASA/NSAIDs use Events	Total	No/occasional use Events	Total	Weight	Odds Ratio M-H, Random, 95% CI
Anderson	10	6012	60	17277	12.4%	0.48 [0.24, 0.93]
Menezes	17	79	108	327	13.4%	0.56 [0.31, 1.00]
Ratnasinghe	43	14838	35	7996	14.8%	0.66 [0.42, 1.03]
Jacobs	37	7769	3455	721041	16.1%	0.99 [0.72, 1.38]
Coogan	18	188	207	2339	14.2%	1.09 [0.66, 1.81]
Schernhammer	37	10292	153	89541	15.7%	2.11 [1.47, 3.02]
Langman	25	48	413	1286	13.4%	2.30 [1.29, 4.10]
Total (95% CI)		**39226**		**839807**	**100.0%**	**1.01 [0.65, 1.55]**
Total events	187		4431			

Heterogeneity: $Tau^2 = 0.28$; $Chi^2 = 35.73$, df = 6 ($P < 0.00001$); $I^2 = 83\%$

Test for overall effect: $Z = 0.04$ ($P = 0.97$)



Capurso G, Schünemann HJ, Terrenato I, Moretti A, Koch M, Muti P, Capurso L, Delle Fave G.
Meta-analysis: the use of non-steroidal anti-inflammatory drugs and pancreatic cancer risk for different exposure categories.
Aliment Pharmacol Ther. 2007 Oct 15;26(8):1089-99.

3. Directness of Evidence

- indirect comparisons
  - interested in A versus B
  - have A versus C and B versus C
  - ASA/XR dipyridamole vs ASA vs clopidogrel
- differences in
  - patients (low versus high risk atrial fibrillation)
  - interventions (all antiplatelets)
  - outcomes (mortality, long-term QoL, short –term functional capacity, laboratory exercise, C-reactive Protein)

What can raise quality?

3 Factors:

- large magnitude of effect can upgrade (RRR 50%)
  - very large two levels (RRR 80%)
  - common criteria
    - everyone used to do badly
    - almost everyone does well
  - Oral anticoagulation for mechanical heart valves
- dose response relation
  (higher INR – increased bleeding)
- all plausible confounding may be working to reduce the demonstrated effect or increase the effect if no effect was observed

Quality assessment criteria

Quality of evidence	Study design	Lower if	Higher if
High	Randomised trial	**Study quality**: Serious limitations Very serious limitations Important **inconsistency** **Directness**: Some uncertainty Major uncertainty **Sparse or imprecise data** High probability of **reporting bias**	**Strong association:** Strong, no plausible confounders Very strong, no major threats to validity Evidence of a **Dose response gradient** All **plausible confounding** may be working to reduce the demonstrated effect or increase the effect if none was observed
Moderate			
Low	Observational study		
Very low			

Evidence Profile

Oseltamivir for treatment of H5N1 infection:

Quality assessment						Summary of findings					
						No of patients		Effect		Quality	Importance
No of studies (Ref)	Design	Limitations	Consistency	Directness	Other considerations	Oseltamivir	Placebo	Relative (95% CI)	Absolute (95% CI)		
Healthy adults:											
Mortality											
0	-	-	-	-	-	-	-	-	-		**9**
Hospitalisation (Hospitalisations from influenza – influenza cases only)											
5 (TJ 06)	Randomised trial	No limitations	-	Major uncertainty (-2)[1]	Imprecise or sparse data (-1)	-	-	OR 0.22 (0.02 to 2.16)	-	⊕○○○ Very low	**6**
Duration of hospitalization											
0	-	-	-	-	-	-	-	-	-	-	**7**
LRTI (Pneumonia - influenza cases only)											
5 (TJ 06)	Randomised trial	No limitations	-	Major uncertainty (-2)[1]	Imprecise or sparse data (-1)[2]	2/982 (0.2%)	9/662 (1.4%)	RR 0.149 (0.03 to 0.69)	-	⊕○○○ Very low	**8**
Duration of disease (Time to alleviation of symptoms/median time to resolution of symptoms – influenza cases only)											
5[3] (TJ 06) (DT 03)	Randomised trials	No limitations[4]	Important inconsistency (-1)[5]	Major uncertainty (-2)[1]	-	-	-	HR 1.30[3] (1.13 to 1.50)	-	⊕○○○ Very low	**5**
Viral shedding (Mean nasal titre of excreted virus at 24h)											
2[6] (TJ 06)	Randomised trials	No limitations	[-7]	Major uncertainty (-2)[1]	None	-	-	-	WMD -0.73[8] (-0.99 to -0.47)	⊕⊕○○ Low	**4**
Outbreak control											
0	-	-	-	-	-	-	-	-	-	-	**4**
Resistance											
0	-	-	-	-	-	-	-	-	-	-	**7**
Serious adverse effects (Mention of significant or serious adverse effects)											
0[9]	-	-	-	-	-	-	-	-	-	-	**7**
Minor adverse effects [10] (number and seriousness of adverse effects)											
3[11] (TJ 06)	Randomised trials	No limitations	[-12]	Some uncertainty (-1)[13]	Imprecise or sparse data (-1)[14]	-	-	OR range[15] (0.56 to 1.80)	-	⊕⊕○○ Low	
Cost of drugs											
0	-	-	-	-	-	-	-	-	-	-	**4**

Oseltamivir for Avian Flu

Summary of findings:

- No clinical trial of oseltamivir for treatment of H5N1 patients.
- 4 systematic reviews and health technology assessments (HTA) reporting on 5 studies of oseltamivir in seasonal influenza.
  - Hospitalization: OR 0.22 (0.02 – 2.16)
  - Pneumonia: OR 0.15 (0.03 - 0.69)
- 3 published case series.
- Many in vitro and animal studies.
- No alternative that is more promising at present.

Strength of recommendation

- “The strength of a recommendation reflects the extent to which we can, across the range of patients for whom the recommendations are intended, be confident that desirable effects of a management strategy outweigh undesirable effects.”
- Strong or weak

Quality of evidence & strength of recommendation

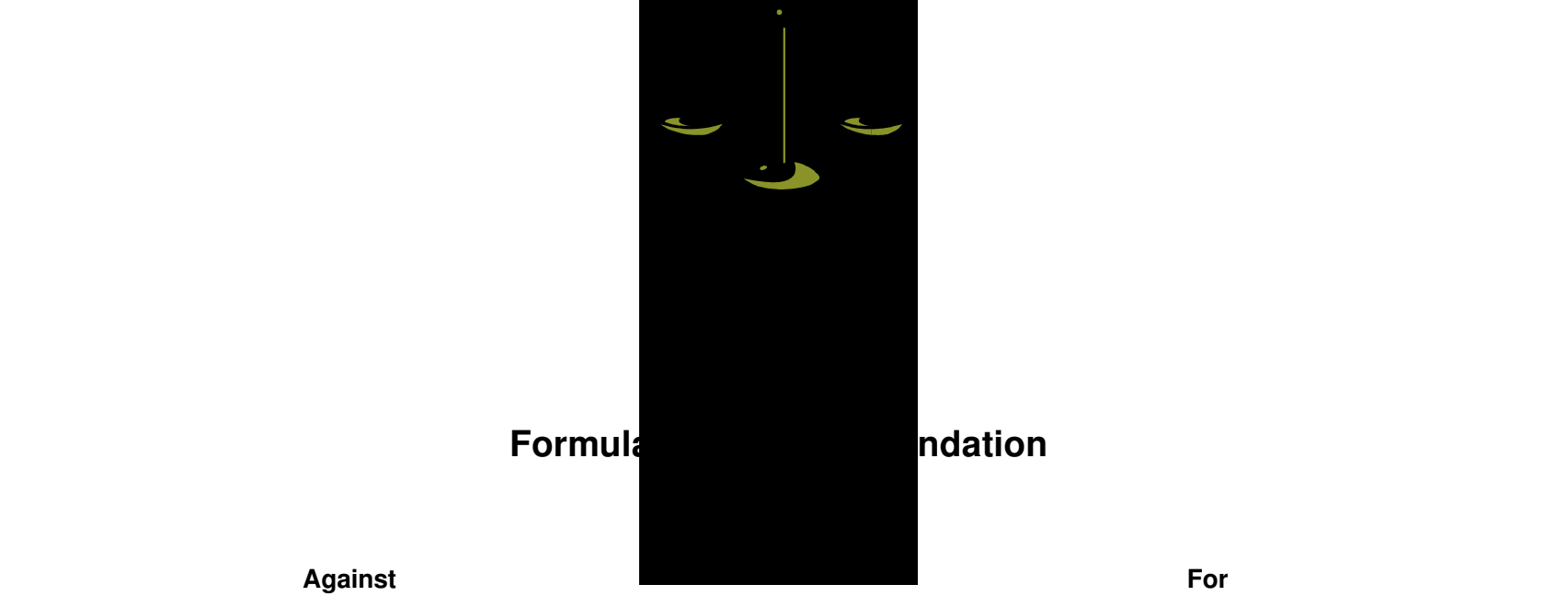
- Linked but no automatism
- Other factors beyond the quality of evidence influence our confidence that adherence to a recommendation causes more benefit than harm
- Systems/approaches failed to make this explicit
- GRADE separates quality of evidence from strength of recommendation

Factors determining strength of recommendation

Factors that can strengthen a recommendation	**Comment**
Quality of the evidence	The higher the quality of evidence, the more likely is a strong recommendation.
Balance between desirable and undesirable effects	The larger the difference between the desirable and undesirable consequences, the more likely a strong recommendation warranted. The smaller the net benefit and the lower the certainty for that benefit, the more likely is a weak recommendation.
Values and preferences	The greater the variability in values and preferences, or uncertainty in values and preferences, the more likely weak recommendation warranted.
Costs (resource allocation)	The higher the costs of an intervention – that is, the more resources consumed – the less likely is a strong recommendation warranted

Evaluating des sirable effects

Desirable << Undesirable Effects | Desirable ?< Undesirabl | ?> Undesirable Effects | Desirable >> Undesirable Effects



Formula ndation

Against | **For**

Strong	Weak	Weak	Strong
↓↓	?↓	↑?	↑↑
1	2	2	1

Content

- Background and rationale for revisiting guideline methodology and GRADE
- GRADE approach
  - Quality of evidence
  - Strength of recommendations
- Examples of Guidelines based on the GRADE approach

Example: Oseltamivir for Avian Flu

Recommendation: In patients with confirmed or strongly suspected infection with avian influenza A (H5N1) virus, clinicians should administer oseltamivir treatment as soon as possible (strong recommendation, very low quality evidence – GRADE 1D).

Values and Preferences

Remarks: This recommendation places a high value on the prevention of death in an illness with a high case fatality. It places relatively low values on adverse reactions, the development of resistance and costs of treatment.

Schunemann et al. The Lancet ID, 2007

ACCP: Acute coronary syndrome

For all patients presenting with NSTE ACS, without a clear allergy to aspirin, we recommend immediate aspirin, 75 to 325 mg po, and then daily, 75 to 162 mg po (strong recommendation, high quality evidence – GRADE 1A).

Conclusion

- clinicians, policy makers need summaries
  - quality of evidence
  - strength of recommendations
- explicit rules
  - transparent, informative
- GRADE
  - four categories of quality of evidence
  - two grades for strength of recommendations
  - transparent, systematic by and across outcomes
  - applicable to diagnosis
  - wide adoption

Disclosure

Relevant Financial Relationships

- Member of the GRADE working group: honoraria related to this work (guideline development) deposited into research accounts
  - Chiesi, AstraZeneca
- No honoraria from pharmaceutical industry policy since April 2008

Off label medication

- None mentioned

The clinical question

Population:	In smokers with COPD
Intervention:	does beta-carotene suppl
Comparison:	compared to no suppl.
Outcomes:	reduce the risk of COPD symptoms, lung cancer and death and improve PFTs?

4. Publication Bias

- publication bias
  - number of small studies

5. Imprecision

- small sample size
  - small number of events
  - wide confidence intervals
  - uncertainty about magnitude of effect

A COPD guidelines

7.6. Mucolytic/antioxidant therapy

These include drugs such as:

- ambroxol
- erdosteine
- carbocysteine
- iodinated glycerol

The regular use of these drugs has been evaluated in a number of studies with little evidence of any effect on lung function.

Data from a Cochrane review of the studies supports a role for these drugs in reducing the number of exacerbations of chronic bronchitis [33].

There is better evidence that *N*-acetylcysteine, a drug with mucolytic and anti-oxidant actions, can reduce the number of exacerbations of COPD and this is currently under study in a large prospective trial [34].

And another COPD guideline

Mucolytic (mucokinetic, mucoregulator) agents (ambroxol, erdosteine, carbocysteine, iodinated glycerol). The regular use of mucolytics in COPD has been evaluated in a number of long-term studies with controversial results[150-152]. Although a few patients with viscous sputum may benefit from mucolytics[153, 154], the overall benefits seem to be very small, and the widespread use of these agents cannot be recommended at present (**Evidence D**).

Antioxidant agents. Antioxidants, in particular N-acetylcysteine, have been reported in small studies to reduce the frequency of exacerbations, leading to speculation that these medications could have a role in the treatment of patients with recurrent exacerbations[155-158] (**Evidence B**). However, a large randomized controlled trial found no effect of N-acetylcysteine on the frequency of exacerbations, except in patients not treated with inhaled glucocorticosteroids[159].

What to do?

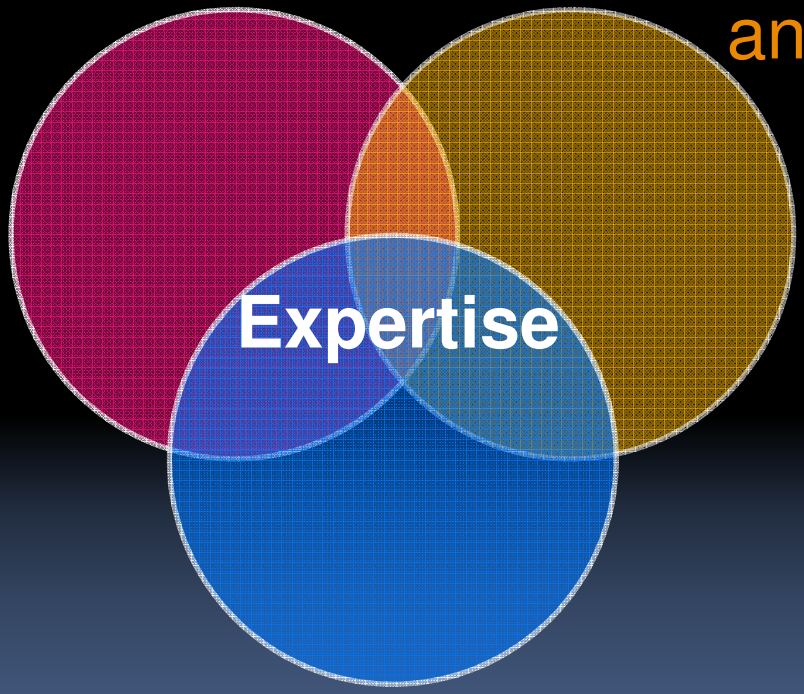
Antibiotics for Pneumonia in COPD and asthma

1) 65 year old man with COPD, terminal lung cancer and chronic pain has come to terms with his condition, has issues in order, said his goodbyes. He wishes to receive palliative care. He develops pneumococcal pneumonia.
2) 85 year old woman with COPD, severely demented, incontinent, and mute, without friends or family and in apparent discomfort. She develops pneumococcal pneumonia.
3) 30 year asthmatic, mother of two and otherwise healthy develops pneumococcal pneumonia.

Evidence based clinical decisions

Clinical state and circumstances

Patient values and preferences



Research evidence

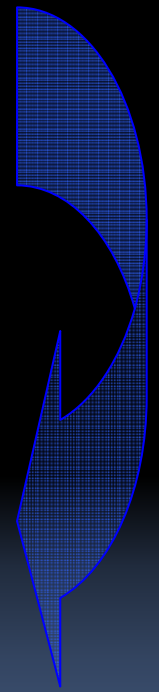
Haynes et al. 2002

Classification schemes

Category of evidence:
Ia—evidence for meta-analysis of randomised controlled trials
Ib—evidence from at least one randomised controlled trial
IIa—evidence from at lease one controlled study without randomisation
IIb—evidence from at lease one other type of quasi-experimental study
III—evidence from non-experimental descriptive studies, such as comparative studies, correlation studies, and case-control studies
IV—evidence from expert committee reports or opinions or clinical experience of respected authorities, or both

Strength of recommendation:
A—directly based on category I evidence
B—directly based on category II evidence or extrapolated recommendation from category I evidence
C—directly based on category III evidence or extrapolated recommendation from category I or II evidence
D—directly based on category IV evidence or extrapolated recommendation from category I, II or III evidence

BMJ VOLUME 318 27 FEBRUARY 1999 www.bmj.com



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, methodologist[i]
- 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, director[n]
- 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[r]
- Holger J Schünemann, professor[g,s]
- 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 Profiles

Quality assessment							Summary of findings					Importance
							No of patients		Effect		Quality	
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Erythropoesis stimulants (epo)	placebo	Relative (95% CI)	Absolute		
All cause mortality (follow-up 4 - 36 months)												
35	randomised trial	no serious limitations	serious[1]	no serious indirectness	serious[2]	none	1008/3825	830/3093	HR 1.11 (1 to 1.22)	25 more per 1000 (from 0 more to 49 more)	⊕⊕OO LOW	CRITICAL
								10%		10 more per 1,000		
								50%		36 more per 1,000		
Thromboembolic events												
30	randomised trial	serious[3]	serious[1]	no serious indirectness	no serious imprecision	none	218/3355	112/2737	RR 1.69 (1.36 to 2.1)	28 more per 1000 (from 15 more to 45 more)	⊕⊕OO LOW	CRITICAL
								1%		6 more per 1,000		
								8%		55 more per 1,000		
Complete response of tumor to chemotherapy												
5	randomised trial	no serious limitations	serious[1]	no serious indirectness	no serious imprecision	reporting bias[4]	216/344	211/344	RR 1.0 (0.92 to 1.1)	0 fewer per 1000 (from 49 fewer to 61 more)	⊕⊕OO LOW	CRITICAL
Transfusion rates (follow-up 4-26 weeks)												
34	randomised trial	no serious limitations	serious[5]	no serious indirectness	no serious imprecision	none	864/2859	1110/2351	RR 0.63 (0.59 to 0.67)	175 fewer per 1000 (from 156 fewer to 194 fewer)	⊕⊕⊕O MODERATE	CRITICAL
								25%		92 fewer per 1,000		
								75%		277 fewer per 1,000		
Increase > 2 mg/dL in Hb (mg/dL) (follow-up 4-20 weeks)												
15	randomised trial	no serious limitations	serious[6]	no serious indirectness	no serious imprecision	strong association[7]	1069/1844	239/1449	RR 3.42 (3.03 to 3.86)	399 more per 1000 (from 335 more to 472 more)	⊕⊕⊕⊕ HIGH	IMPORTANT

[1] Overall heterogeneity not significant, but underlying clinical heterogeneity due to risk of VTE, treatment regimens, and epo potocols (atarting and stopping Hb).

[2] CI includes no effect and clinically improtant increase in mortality

[3] Criteria for determining and reporting VTE variable in studies; trials reporting varying combinations of DVT, PE, TIA, stroke, and MI

[4] Only 5 trials reported this outcome; does not include the largest trials powered for mortaltiy benefit.

[5] Tests of heterogeneity I square were significant. Reduced risk of transfusion evidence in subgroups defined by different starting Hb level, but size of benefit differes. Clinical heterogeneity in control rate transfusions, tumor type and chemo regimen, and prtocols for determining transusion need.

[6] All trials support substantial benefit but significant heterogeneity in magnitude of benefit; clinical heterogeneity in starting Hb levels, underlying chemo regimens and tumor types, and risk of anemia

[7] Size of RR (3.4 pooled, range 2 to 9) would qualify as large effect

Summary of Findings Tables

Erythropoesis stimulants (epo) compared to placebo for anemia from cancer chemotherapy

Patient or population: patients with anemia from cancer chemotherapy
Settings: Outpatient cancer treatment
Intervention: Erythropoesis stimulants (epo)
Comparision: placebo

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk placebo	Corresponding risk Erythropoesis stimulants (epo)	Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
All cause mortality (follow-up: 4 - 36 months)	**Population**		**HR 1.11** (1 to 1.22)	6918 (35)	⊕⊕OO **low**[1,2]	
	268 per 1000	**293 per 1000** (268 to 317)				
	Low risk population					
	100 per 1000	**110 per 1000** (100 to 121)				
	High risk population					
	500 per 1000	**537 per 1000** (500 to 571)				
Thromboembolic events	**Population**		**RR 1.69** (1.36 to 2.1)	6092 (30)	⊕⊕OO **low**[1,3]	
	41 per 1000	**69 per 1000** (56 to 86)				
	Low risk population					
	10 per 1000	**17 per 1000** (14 to 21)				
	High risk population					
	80 per 1000	**135 per 1000** (109 to 168)				
Complete response of tumor to chemotherapy	**613 per 1000**	**613 per 1000** (564 to 674)	**RR 1.0** (0.92 to 1.1)	688 (5)	⊕⊕OO **low**[1,4]	

[1] Overall heterogeneity not significant, but underlying clinical heterogeneity due to risk of VTE, treatment regimens, and epo potocols (atarting and stopping Hb).

[2] CI includes no effect and clinically improtant increase in mortality

[3] Criteria for determining and reporting VTE variable in studies; trials reporting varying combinations of DVT, PE, TIA, stroke, and MI

[4] Only 5 trials reported this outcome; does not include the largest trials powered for mortaltiy benefit.

[5] Tests of heterogeneity I square were significant. Reduced risk of transfusion evidence in subgroups defined by different starting Hb level, but size of benefit differes. Clinical heterogeneity in control rate transfusions, tumor type and chemo regimen, and prtocols for determining transusion need.

[6] All trials support substantial benefit but significant heterogeneity in magnitude of benefit; clinical heterogeneity in starting Hb levels, underlying chemo regimens and tumor types, and risk of anemia

[7] Size of RR (3.4 pooled, range 2 to 9) would qualify as large effect

Strength of recommendation

- “The strength of a recommendation reflects the extent to which we can, across the range of patients for whom the recommendations are intended, be confident that desirable effects of a management strategy outweigh undesirable effects.”

Desirable and undesirable effects

- Desirable effects
  - Mortality
  - improvement in quality of life, fewer hospitalizations/infections
  - reduction in the burden of treatment
  - reduced resource expenditure
- Undesirable effects
  - deleterious impact on morbidity, mortality or quality of life, increased resource expenditure

Determinants of the strength of recommendation

Factors that can strengthen a recommendation	Comment
Quality of the evidence	
Balance between desirable and undesirable effects	
Values and preferences	
Costs (resource allocation)	

Determinants of the strength of recommendation

Factors that can strengthen a recommendation	Comment
Quality of the evidence	The higher the quality of evidence, the more likely is a strong recommendation.
Balance between desirable and undesirable effects	The larger the difference between the desirable and undesirable consequences, the more likely a strong recommendation warranted. The smaller the net benefit and the lower certainty for that benefit, the more likely weak recommendation warranted.
Values and preferences	The greater the variability in values and preferences, or uncertainty in values and preferences, the more likely weak recommendation warranted.
Costs (resource allocation)	The higher the costs of an intervention – that is, the more resources consumed – the less likely is a strong recommendation warranted

Determinants of the strength of recommendation

Factors that can strengthen a recommendation	Comment
Quality of the evidence	Low
Balance between desirable and undesirable effects	Moderately balance
Values and preferences	Fair bit of variability
Costs (resource allocation)	Relatively high cost

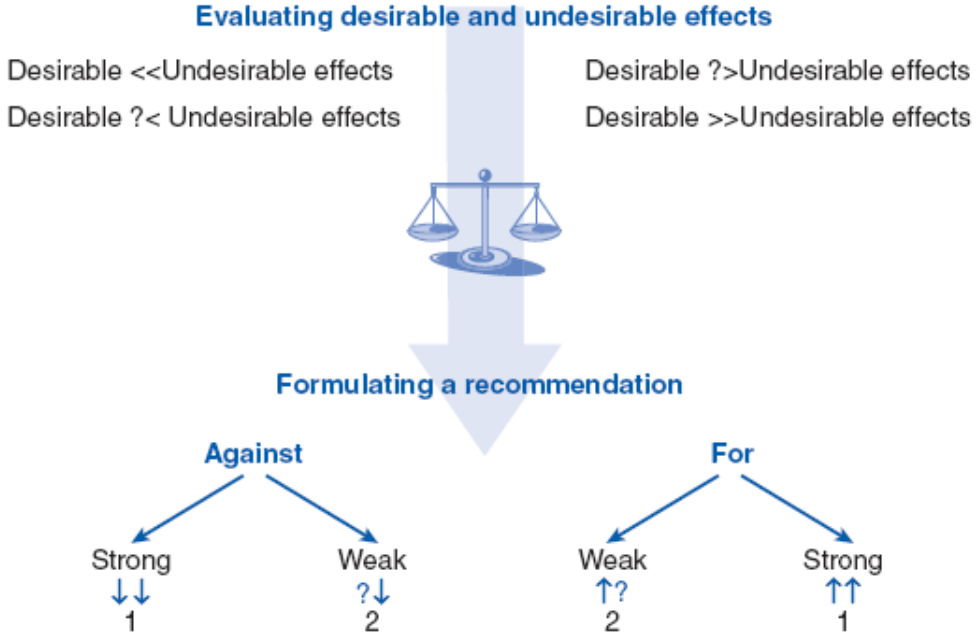
Determinants of the strength of recommendation

Factors that can weaken the strength of a recommendation. Example:	Decision	Explanation
Lower quality evidence	□ Yes □ No	
Uncertainty about the balance of benefits versus harms and burdens	□ Yes □ No	
Uncertainty or differences in values	□ Yes □ No	
Uncertainty about whether the net benefits are worth the costs	□ Yes □ No	

Table. Decisions about the strength of a recommendation
Frequent “yes” answers will increase the likelihood of a weak recommendation

Developing recommendations

Strength of Recommendations



The figure describes the balance between important benefits and downsides relate to a recommendation. The process begins by evaluating whether desirable effects outweigh undesirable effects or vice versa. Moving on to making a recommendation requires a decision: if the balance is clear, a strong recommendation for or against an action follows (<< and >> denote a clear balance). If the balance is not clear, a weak recommendation for or against an action follows (?< and ?> denote a balance that is not clear). Widely differing values (the importance or preference patients assign to a certainhealth state) can also lead to a less clear balance of benefits versus downsides.