

# Rating quality evidence, grading strength recommendation

- informative summaries for consumer
- eminently useful, but proliferation
  - Australian National and MRC
  - Oxford Center for Evidence-based Medicine
  - Scottish Intercollegiate Guidelines (SIGN)
  - US Preventative Services Task Force
  - American professional organizations
    - AHA/ACC, ACCP, AAP, Endocrine society, etc....
- cause of confusion, dismay

# A common international grading system?

- international group
  - methodologists, guideline developers
  - Australian NMRC, SIGN, USPSTF, WHO, NICE, Oxford CEBM, CDC, CC
- GRADE (*Grades of recommendation, assessment, development and evaluation*)
- ~ 25 meetings over last 10 years
  - (~10 - 60 attendants)

# GRADE Uptake

Agencia sanitaria regionale, Bologna, Italia  
Agency for Health Care Research and Quality (AHRQ)  
Allergic Rhinitis and Group - Independent Expert Panel  
American College of Cardiology Foundation  
American College of Chest Physicians  
American College of Emergency Physicians  
**American College of Physicians**  
American Endocrine Society  
American Society of Gastrointestinal Endoscopy  
American society of Interventional Pain Physicians  
**American Thoracic Society (ATS)**  
**BMJ Clinical Evidence**  
**British Medical Journal**  
Canadian Agency for Drugs and Technology in Health  
Centers for Disease Control  
**Cochrane Collaboration**  
EBM Guidelines Finland  
Emergency Medical Services for Children National Resource Center  
European Association for the Study of the Liver  
European Respiratory Society  
European Society of Thoracic Surgeons  
Evidence-based Nursing Sudtiro, Alta Adiga, Italy  
Finnish Office of Health Technology Assessment  
German Agency for Quality in Medicine

Infectious Disease Society of America  
Japanese Society of Oral and Maxillofacial Radiology  
Joslin Diabetes Center  
Journal of Infection in Developing Countries  
Kidney Disease International Guidelines Organization  
National and Gulf Centre for Evidence-based Medicine  
**National Institute for Clinical Excellence (NICE)**  
National Kidney Foundation  
Norwegian Knowledge Centre for the Health Services  
Ontario MOH Medical Advisory Secretariat  
Panama and Costa Rica National Clinical Guidelines Program  
Polish Institute for EBM  
**Scottish Intercollegiate Guideline Network (SIGN)**  
Society of Critical Care Medicine  
Society of Pediatric Endocrinology  
Society of Vascular Surgery  
Spanish Society of Family Practice (SEMFYC)  
Stop TB Diagnostic Working Group  
Surviving sepsis campaign  
Swedish Council on Technology Assessment in Health Care  
Swedish National Board of Health and Welfare  
University of Pennsylvania Health System for EB Practice  
**UpToDate**  
**World Health Organization (WHO)**

# What are we grading?

- two components
- quality of body of evidence
  - extent to which confidence in estimate of effect adequate to support decision
    - high, moderate, low, very low
- strength of recommendation
  - strong and weak

# Quality assessment criteria

Study Design	Quality of Evidence	Lower if	Higher if
Randomised trial →	High	Risk of bias -1 Serious -2 Very serious	Large effect +1 Large +2 Very large
	Moderate	Inconsistency -1 Serious -2 Very serious	Dose response +1 Evidence of a gradient
Observational study →	Low	Indirectness -1 Serious -2 Very serious	All plausible confounding +1 Would reduce a demonstrated effect or
	Very low	Imprecision -1 Serious -2 Very serious	+1 Would suggest a spurious effect when results show no effect
		Publication bias -1 Likely -2 Very likely	

## Beta blockers in non-cardiac surgery

Quality Assessment							Summary of Findings		
							Quality	Relative Effect (95% CI)	Absolute risk difference
Outcome	Number of participants (studies)	Serious Risk of Bias	Consistency	Directness	Precision	Reporting Bias			
Myocardial infarction	10,125 (9)	No	OK	OK	OK	Undetected	High	0.71 (0.57 to 0.86)	1.5% fewer (0.7% fewer to 2.1% fewer)
Mortality	10,205 (7)	No	Possible ↓	OK	Imprecise	Undetected	Moderate or low	1.23 (0.98 – 1.55)	0.5% more (0.1% fewer to 1.3% more)
Stroke	10,889 (5)	No	OK	OK	OK	Undetected	High	2.21 (1.37 – 3.55)	0.5% more (0.2% more to 1.3% more)

# Resource use: just another outcome?

- yes and no
- who benefits?
  - different payers bear costs across societies and within (age)
- costs vary much more than other outcomes
  - across/within jurisdictions, over time
- even when resource use same implications differ
  - year's supply expensive drug
  - nurses' salary in U.S., 6 in Poland, 30 in China
- unbearable lightness of costs
- may decide to omit from consideration

# GRADE's approach to resource use

- identify viewpoint
- identify important resource use items
- find relevant evidence
- evaluate evidence quality
  - may differ across resource use items
  - RCTs start high, observational low
  - 5 categories for rating down, 3 up
  - economic analysis criteria inapplicable
- value resources in terms of cost



# GRADE approach evidence quality

- reasons for risk of bias
  - failure ITT (methadone vs buprenorphine)
  - reliance on patient recall
  - reliance on imputation
- directness often major issue
  - older studies - different practice patterns
  - only cost provided
  - modeling necessary

# Evidence summary

- as other outcomes, need systematic review
- quality of evidence, summary of findings
  - "balance sheet", special form evidence profile
- resource use and not just costs
  - can judge whether resource use applicable to local setting
  - focus on items relevant to them (pharmacy)
  - apply unit costs to local setting

# Example question

- patients
  - women with pre-eclampsia
- intervention
  - intravenous magnesium
- RCT done in 33 countries
  - over 9,000 patients
- health system perspective

**Table 2 | Summary of findings on whether clinicians should use magnesium sulphate to prevent eclampsia: resource use viewed from the perspective of the health system**

Resource	Cost*	Typical absolute effect (95% CI)	No of participants (studies)	Quality of evidence	Comment
<b>Magnesium sulphate ampoules (6×10 ml ampoules/patient)</b>					
Setting:					
High income countries	\$20 more/patient		9996	High	
Middle income countries	\$3 more/patient				
Low income countries	\$5 more/patient				
<b>Administration of magnesium sulphate (1 ampoule/patient)</b>					
Setting:					
High income countries	\$66/patient		9996	High	Resources for giving magnesium sulphate included midwives' time (main cost), intravenous cannula or needles, syringes, intravenous fluids, and the drug
Middle income countries	\$14/patient				
Low income countries	\$8/patient				
<b>Other hospital resources (varied widely)</b>					
Setting:					
High income countries	\$12 839	\$20 less/ patient (\$0 to \$60)	9.996	Moderate†	Use of other hospital resources varied greatly in both intervention and control groups. Other hospital costs have been adjusted for on the basis of the influence of eclampsia to control for the many other factors that influenced these costs
Middle income countries	\$1 416	\$4 less/ patient (\$0 to \$10)			
Low income countries	\$157	\$2 less/ patient (\$1 to \$3)			

\* $\$1 = \text{£}0.5 = \text{€}0.7$ .

†Evidence comes from randomised trials and there was no reason to grade down for study limitations, imprecision, inconsistency, indirectness, or publication bias.

‡The confidence interval was wide so the evidence was graded down for imprecision.

# Example question 2

- patients
  - opioid dependent
- intervention
  - buprenorphine versus methadone
- 2 RCTs
- societal perspective

**Table 1 - Example of resource use evidence profile<sup>o</sup>**

**Question:** Should Buprenorphine maintenance flexible doses vs Methadone maintenance flexible doses be used for opioid maintenance treatment?

**Patient or population:** Opiate dependents; **Setting:** Outpatients in USA, Australia, Austria, Switzerland, UK; **Viewpoint:** societal

Studies (follow up)	Quality assessment					No of patients	Summary of resources and costs		Overall Quality
	Design	Limitations	Inconsistency	Indirectness	Imprecision		Resources Costs per patient (1999 AU \$)		
Drugs							Methadone	Buprenorphine	
Harris 2005 (1 year) <sup>a</sup>	RCT	Serious limitations <sup>o</sup>	No	Some uncertainty <sup>g</sup>	Small sample size	139	Resources (mean daily)		Very low
							50 mg	14 mg	
							Costs (annual)		
							1,122 (85 SE)	1,785 (204 SE)	
Doran 2003 (6 months) <sup>a</sup>	RCT	No	No	Some uncertainty <sup>g</sup>		405	Resources (mean daily)		Moderate
							57 mg	11 mg	
							Costs (6 months)		
							37 (33 SD)	459 (461 SD)	
<b>Other healthcare costs</b>									
Harris 2005 (1 year) <sup>b</sup>	RCT	Serious limitations <sup>o</sup>	No	Some uncertainty <sup>g</sup>	Small sample size	139	Resources		Very low
							NA	NA	
							Costs (annual)		
							2,500 (489 SE)	3,316 (667 SE)	
Doran 2003 (6 months) <sup>c</sup>	RCT	No	No	Some uncertainty <sup>g</sup>		405	Resources		Moderate
							NA	NA	
							Costs (6 months)		
							1,378 (NA)	1,270 (NA)	
<b>Crime costs</b>									
Harris 2005 (1 year) <sup>d</sup>	RCT	Very serious limitations <sup>f</sup>	No	Some uncertainty <sup>g</sup>	Small sample size	139	Resources		Very low
							NA	NA	
							Costs (annual) <sup>h</sup>		
							13,223 (10,209 SE)	6,265 (2,028 SE)	

<sup>o</sup> In this example we decided not to pool resource data from different studies due to insufficient information provided

NA = not available SD = Standard deviation SE = Standard Error

<sup>a</sup> Including dispensing fee

<sup>b</sup> Includes other prescription and OTC drugs, prescriber, inpatient, outpatient, emergency, ambulance, counseling, allied health and pathology services

<sup>c</sup> Include staff time (i.e. face-to-face contact and preparation time), diagnostic procedures and facility level (Supplies, consumables, capital, equipment, ancillary support including administration, management, security, etc.

<sup>d</sup> Healthcare costs from assault, loss of income by the victims of crime, depreciated value of property damaged, stolen or obtained fraudulently, detection, prosecution and imprisonment

<sup>e</sup> Some limitations because of incomplete outcome data

<sup>f</sup> Some limitations because of incomplete outcome data and crucial limitations for self reported crime data

<sup>g</sup> All the studies were conducted within the Australia health system (while the recommendation was global)

<sup>h</sup> Doses for methadone and buprenorphine derived from Mattick 2003 study, at the 10<sup>th</sup> week.

<sup>i</sup> The average cost of crime was substantial across the sample by these reported costs were associated with just a few participants. ) 90% of the sample randomized to methadone and 96% of that randomized on buprenorphine reported non-involvement in property crime. Indeed the majority of patients reported non criminal activity during the trial (6/66 patients for methadone and 3/73 for buprenorphine). (page 86)

# Strength of recommendations

- degree of confidence that desirable effects of adhering to recommendation outweigh undesirable effects.



- strong recommendation
  - benefits clearly outweigh risks/hassle/cost
  - risk/hassle/cost clearly outweighs benefit

## Determinants of strength of recommendation

Factor	Comment
Balance between desirable and undesirable effects	The larger the difference between the desirable and undesirable effects, the higher the likelihood that a strong recommendation is warranted. The narrower the gradient, the higher the likelihood that a weak recommendation is warranted
Quality of evidence	The higher the quality of evidence, the higher the likelihood that a strong recommendation is warranted
Values and preferences	The more values and preferences vary, or the greater the uncertainty in values and preferences, the higher the likelihood that a weak recommendation is warranted
Costs (resource allocation)	The higher the costs of an intervention—that is, the greater the resources consumed—the lower the likelihood that a strong recommendation is warranted



# Significance of strong vs weak

- variability in patient preference
  - strong, almost all same choice (> 90%)
  - weak, choice varies appreciably
- interaction with patient
  - strong, just inform patient
  - weak, ensure choice reflects values
- use of decision aid
  - strong, don't bother
  - weak, use the aid
- quality of care criterion
  - strong, consider
  - weak, don't consider

# Value and preference statements

- underlying values and preferences always present
- sometimes crucial
- important to make explicit

# Values and preferences

Stroke guideline: patients with TIA  
clopidogrel over aspirin (Grade 2B).

*Underlying values and preferences:* This recommendation to use clopidogrel over aspirin places a relatively high value on a small absolute risk reduction in stroke rates, and a relatively low value on minimizing drug expenditures.

# Values and preferences

peripheral vascular disease: aspirin be used instead of clopidogrel (Grade 2A).

*Underlying values and preferences:*

This recommendation places a relatively high value on avoiding large expenditures to achieve small reductions in vascular events.

# Summary

- GRADE provide transparent structural framework for developing and presenting recommendations
- increasingly widely adopted
- refinements required, provides framework for dealing with resource use/cost