How to make transparent judgments about the quality of evidence

Interactive Workshop
GIN 2013 San Francisco

McMaster University
Inspiring Innovation and Discovery

DUTCH COCHRANE CENTRE

Facilitators
Holger Schünemann
Nancy Santesso
Reem Mustafa
Miranda Langendam

Outline

Aim of today: how to write transparent footnotes and comments in GRADE evidence profiles and Summary of Findings Tables

- Background: rationale for this workshop
- GRADE approach – focus on footnotes & comments

} 20 min

- Hands-on work: (60 min)
  - Writing footnotes on selected examples (work in pairs and large group discussion)
  - Examples of footnotes: make a judgment on usefulness and interpretation (small group work)
- Wrap up (10 min)

Footnotes explain choices and judgments made in a GRADE evidence profile

Question: Should oseltamivir vs. no antiviral treatment be used for influenza?

Quality assessment							Summary of Findings				
Participants (studies)	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With no antiviral treatment	With oseltamivir		Risk with no antiviral treatment	Absolute effect with Oseltamivir (95% CI)
Mortality											
681 (3 studies)	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	undetected[1]	LOW[1]	59/242 (24.4%)	31/439 (7.1%)	adj OR 0.23 (0.13 to 0.43)	240 deaths per 1000	172 fewer deaths per 1000 (from 120 to 201 fewer)
Hospitalisation											
150710 (5 studies)	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	undetected[4]	LOW[4]	1238/100585 (1.2%)	431/50125 (0.86%)	adj OR 0.75 (0.66 to 0.89)	12 hospitalisations per 1000	3 fewer hospitalisations per 1000 (from 1 to 4 fewer)
ICU admissions/mechanical ventilation/respiratory failure											
1032 (6 studies[5])	Serious[5]	serious[6]	no serious indirectness	no serious imprecision	undetected[1]	VERY LOW[1,6] due to risk of bias, inconsistency	-	200/1032 (19.4%)	-	-	
Complications - Pneumonia											
150466 (3 studies)	no serious risk of bias	serious[6]	no serious indirectness	no serious imprecision	undetected[4]	VERY LOW[4,6] due to inconsistency	2111/ 100449 (2.1%)	647/50017 (1.3%)	adj OR 0.83 (0.59 to 1.16)	21 pneumonias per 1000	4 fewer pneumonias per 1000 (from 9 fewer to 3 more)

Hsu J, Santesso N, Mustafa R, et al.
Antivirals for treatment of influenza: a systematic review and meta-analysis of observational studies.
Ann Intern Med 2012, 512-24

GRADE evidence profile

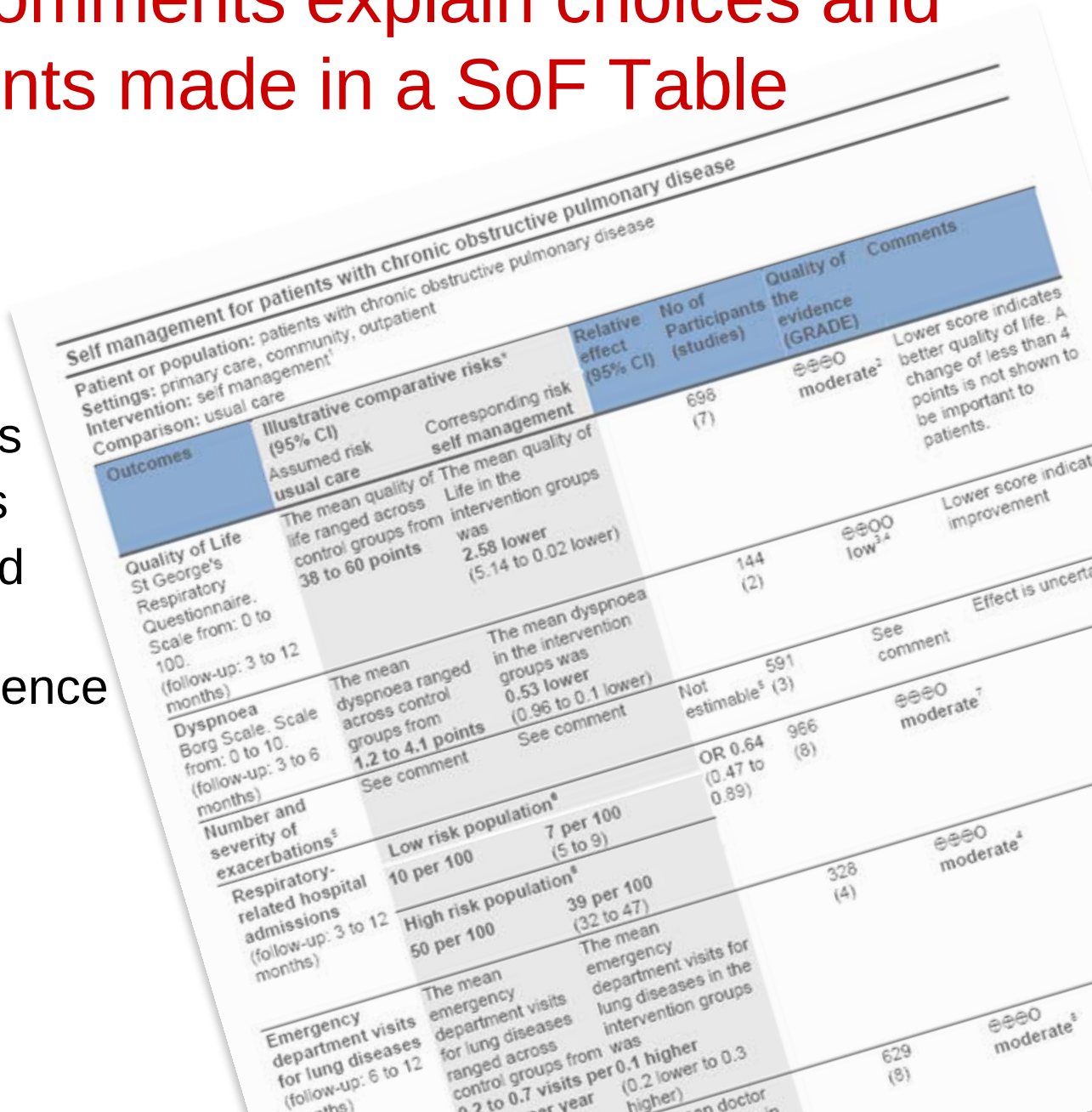
Author(s): Elie Akl & Holger Schunemann **Date:** 2008-09-11
Question: Should parenteral anticoagulation be used in prolonging survival of patients with cancer? **Settings:** Outpatient
Bibliography: EA Akl, FF van Doormaal, M Barba, G Kamath, SY Kim, S Kuipers, S Middeldorp, V Yosuico, H Dickinson, HJ Schünemann. Parenteral anticoagulation for prolonging survival in patients with cancer who have no other indication for anticoagulation. CDSR Reviews. 2007 Issue 3

Quality assessment							Summary of findings					
							No of patients		Effect		Quality	Importance
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	anticoagulation	control	Relative (95% CI)	Absolute		
Survival at 12 months (study follow up)												
5	randomised trials	no serious limitations[1]	no serious inconsistency	no serious indirectness[2]	no serious imprecision	none	339/586 (57.8%)	390/588 (60%)	RR 0.87 (0.8 to 0.95)	78 fewer per 1000 (from 30 to 120 fewer)	⊕⊕⊕⊕ HIGH	CRITICAL
Survival (overall - study follow up at 24 to 84 months)												
5	randomised trials	no serious limitations[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	477/586 (81.4%)	520/588 (85%)	HR 0.77 (0.65 to 0.91)	82 fewer per 1000 (from 28 to 141 fewer)	⊕⊕⊕⊕ HIGH	CRITICAL
DVT												
2	randomised trials	no serious limitations[1]	no serious inconsistency	no serious indirectness	very serious[3]	reporting bias[4]	1/232 (0.4%)	2/226 (4%)	RR 0.61 (0.08 to 4.91)	16 fewer per 1000 (from 37 fewer to 156 more)	⊕OOO VERY LOW	CRITICAL
Major bleeding												
3	randomised trials	no serious limitations[1]	no serious inconsistency	no serious indirectness	serious[3]	reporting bias[5]	8/406 (2%)	6/408 (1.5%)	RR 1.50 (0.26 to 8.8)	7 more per 1000 (from 11 fewer to 117 more)	⊕⊕OO LOW	CRITICAL
Minor bleeding												
3	randomised trials	no serious limitations[1]	no serious inconsistency	no serious indirectness	serious[3]	reporting bias[5]	14/380 (3.7%)	5/380 (1.3%)	RR 2.07 (0.78 to 5.51)	14 more per 1000 (from 3 fewer to 59 more)	⊕⊕OO LOW	IMPORTANT

[1] Unclear concealment in one of the five trials did not lead to downgrading the quality of evidence.
[2] The studies used different LMWHs but indirectness is not likely given the similiarity in results across studies.
[3] The 95% CI includes both negligible effect and appreciable benefit or appreciable harm
[4] Out of 5 included studies, only 2 reported DVT. We assumed that this was based on selective reporting of outcomes. The authors of the study did not provide further information.
[5] Out of 5 included studies, only 3 reported major bleeding. We assumed that this was based on selective reporting of outcomes. The authors of the study did not provide further information.

Footnotes/comments explain choices and judgments made in a SoF Table

- PICO
- Outcomes
- Results
  - Absolute effects
  - Relative effects
  - Participants and studies
- Quality of the Evidence



Self management for patients with chronic obstructive pulmonary disease

Patient or population: patients with chronic obstructive pulmonary disease
Settings: primary care, community, outpatient
Intervention: self management[1]
Comparison: usual care

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk usual care	Corresponding risk self management	Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Quality of Life St George's Respiratory Questionnaire. Scale from: 0 to 100. (follow-up: 3 to 12 months)	The mean quality of life ranged across control groups from 38 to 60 points	The mean quality of Life in the intervention groups was 2.58 lower (5.14 to 0.02 lower)		698 (7)	⊕⊕⊕O moderate[2]	Lower score indicates better quality of life. A change of less than 4 points is not shown to be important to patients.
Dyspnoea Borg Scale. Scale from: 0 to 10. (follow-up: 3 to 6 months)	The mean dyspnoea ranged across control groups from 1.2 to 4.1 points	The mean dyspnoea in the intervention groups was 0.53 lower (0.96 to 0.1 lower)		144 (2)	⊕⊕OO low[3,4]	Lower score indicat improvement
Number and severity of exacerbations[5]	See comment	See comment	Not estimable[5]	591 (3)	See comment	Effect is uncerta
Respiratory-related hospital admissions (follow-up: 3 to 12 months)	Low risk population[6] 10 per 100 High risk population[6] 50 per 100	7 per 100 (5 to 9) 39 per 100 (32 to 47)	OR 0.64 (0.47 to 0.89)	966 (8)	⊕⊕⊕O moderate[7]	
Emergency department visits for lung diseases (follow-up: 6 to 12	The mean emergency department visits for lung diseases ranged across control groups from 0.2 to 0.7 visits per year	The mean emergency department visits for lung diseases in the intervention groups was 0.1 higher (0.2 lower to 0.3 higher)		328 (4)	⊕⊕⊕O moderate[4]	
				629 (8)	⊕⊕⊕O moderate[8]	

Why a workshop on footnotes and comments in GRADE profiles?

- Cochrane Methods Innovation Fund project on evaluation of Summary of Findings tables

- Aim 2: provide guidance and options for footnotes & comments to enhance their usefulness and interpretation
- Methods: extract and analyze footnotes & comments in GRADE evidence profiles in guidelines and SoF Tables in Cochrane reviews

- Team Aim 2: Miranda Langendam, Nancy Santesso, Reem Mustafa, Pauline Heus, Romina Brignardello Petersen, Matt Ventresca, Alonso Carrasco & Holger Schünemann

Methods and Results

- SoF tables in 502 Cochrane reviews (March, 2012)
- Assigned themes to each footnote/comment:
  - down/upgrading quality of evidence, baseline risk, outcome, effect measure, no of studies, study design, results
  - Single RCT, no meta-analysis

- Observed: large variability in formulation, level of detail and clarity, some misconceptions

- Translating our observations in guidance: workshop

Formulate question

Select outcomes

Rate importance

Outcomes across studies

Create evidence profile with GRADEpro

Rate quality of evidence for each outcome

P
I
C
O

Outcome Critical

Outcome Critical

Outcome Important

Outcome Not important

Summary of findings & estimate of effect for each outcome

High ⊕⊕⊕⊕

Moderate ⊕⊕⊕◯

Low ⊕⊕◯◯

Very low ⊕◯◯◯

Grade down

1. Risk of bias
2. Inconsistency
3. Indirectness
4. Imprecision
5. Publication bias

Grade up

1. Large effect
2. Dose response
3. Opposing bias & Confounders

Randomization increases initial quality

Systematic review

Input?

Recommendation or health care action

Grade recommendations

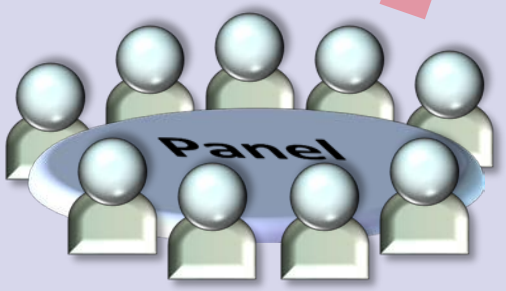
- For or against (direction) ↓↑
- Strong or conditional/weak (strength)

By considering balance of:

- ❑ Quality of evidence
- ❑ Balance benefits/harms
- ❑ Values and preferences

Revise if necessary by considering:

- ❑ Resource use (cost)



Guideline

Grade overall quality of evidence across outcomes based on lowest quality of ***critical*** outcomes

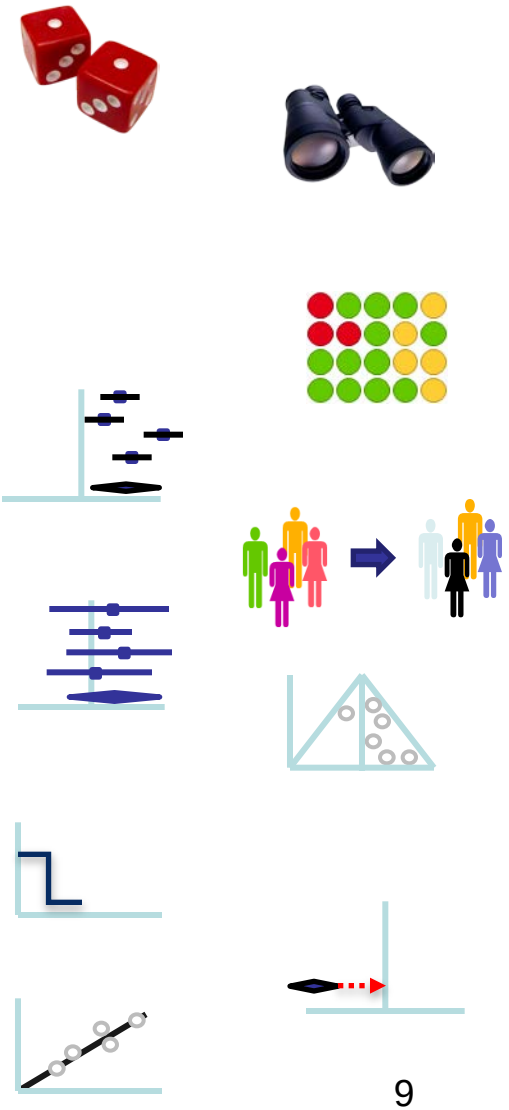
Formulate Recommendations (↓↑ | ⊕...)

- "The panel recommends that ....should..." (↑↑ | ⊕...)
- "The panel suggests that ....should..." (↑? | ⊕...)
- "The panel suggests to **not** ..." (↓? | ⊕...)
- "The panel recommends to **not**..." (↓↓ | ⊕...)

Determinants of quality of a body of evidence

- **RCTs ⊕⊕⊕⊕**
- **observational studies ⊕⊕◯◯**

- **5 factors that can lower quality**
  1. limitations in detailed study design and execution *(risk of bias criteria)*
  2. Inconsistency *(or heterogeneity)*
  3. Indirectness *(PICO and applicability)*
  4. Imprecision
  5. Publication bias
- **3 factors can increase quality**
  1. large magnitude of effect
  2. opposing plausible residual bias or confounding
  3. dose-response gradient



GRADEpro [Overview medication for chronic non-sepcific LBP.grd] ver 3.2.2.20090316

File Add View Options Help

New Open Print Save Undo all changes | Add profile group Add profile Add outcome | Import from RevMan Preview SoF table

Profiles tree

- NSAIDs for chronic low back pain
  - NSAIDs vs placebo for chronic LBP
    - Change in pain intensity
    - Side effects (proportion)
  - Muscle relaxants
  - Antidepressants
  - Opioids
  - Injections

Edit

Outcome: Change in pain intensity — dichotomous / continuous — pooled — Importance: 9 CRITICAL

No of studies: 4

Study design: randomised trials — Quality of evidence: HIGH

Decrease quality of evidence		Increase quality of evidence	
Limitations in design	no	Large effect	no
Inconsistency	no	Plausible confounding would change the effect	no
Indirectness	no	Dose-response gradient	no
Imprecision	no		
Publication bias	unlikely		

Delete Revert Go to Summary of findings

Profile: NSAIDs vs placebo for chronic LBP

Change in pain intensity (follow-up <=12 weeks; measured with: VAS; range of scores: 0-100; Better indicated by lower values) | 4 studies

Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Importance
randomised trials	no serious limitations[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	CRITICAL

Patients (NSAIDs)	Control (placebo)	Relative effect	Absolute effect	Quality
512	508	-	MD 12.40 lower (15.53 to 9.26 lower)	⊕⊕⊕⊕ HIGH

Side effects (proportion) (follow-up <=12 weeks) | 4 studies

Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Importance
randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	IMPORTANT

Patients (NSAIDs)	Control (placebo)	Relative effect	Absolute effect	Quality
242/519 (46.6%)	24.4%	RR 1.24 (1.07 to 1.43)	59 more per 1000 (from 17 more to	⊕⊕⊕O

Footnotes

1. Limitations regarding randomization, withdrawal, co-interventions, follow-up

Add new Change order Edit Delete

Footnotes guidance - general

- Explain source for baseline risk
- Explanations on QoE in footnote, not in comment
- For every factor, report issues with evidence

  “The 95% CI includes both negligible effect and appreciable harm”

- Explain if you decided not to downgrade

  “Unclear concealment in 1 of the 5 trials did not lead to downgrading”

  “Studies used different LMWHs but indirectness not likely given similarity in results across studies”

- No issues, no footnote

Footnotes guidance - factors

- Study limitations
  - No of studies and high Risk of Bias items
- Inconsistency
  - Judgment based on visual inspection CIs, I^2, Chi^2-test?
- Imprecision
  - Judgment based on CI, OIS, clinical thresholds?
- Indirectness
  - Which PICO element causes indirectness?
- Publication bias
  - Judgment based on funnel plot, test, other?

Footnotes guidance - factors

- Magnitude of effect
- Dose-response gradient
- Opposing plausible residual bias or confounding
  - Explain direction

Hands-on work
Writing footnotes on selected examples

- Work in pairs
- Make judgment on quality of evidence and formulate a footnote
- Note discussion points
- Large group discussion

Hands-on work
Make judgement on informativeness

- Work in small groups + tutor
- Use the examples, focus on footnotes/comments
- Discuss:
  - What is your overall opinion of the explanations (footnotes and comments)?
  - Do you find them transparent?
  - How could the explanations be improved?
    - e.g. leaving out information, adding information