**Holger Schünemann**
Professor and Chair, Dept. of Clinical Epidemiology & Biostatistics
Professor of Medicine
Michael Gent Chair in Healthcare Research
McMaster University, Hamilton, Canada

**Rome, March 1, 2013**

# GRADE

## Acknowledgment

GRADE Working Group

# Disclosure

- No direct/personal for-profit payments to me or my research group
- Co-chair of GRADE working group
- Cochrane Collaboration
  - Co-convenor of the Applicability and Recommendations Methods Group
  - Various other functions
- IQWiG Scientific Board

# Content

What is GRADE

- Quality of evidence (certainty or confidence in effects)
- Recommendations (framework for developing)

# A systematic review of randomized trials in cancer patients

**Table 1. Summary of Findings Table Showing the Relative Risks and Absolute Effects over 12 Months for Each Important Outcome after Treatment with a Low-Molecular-Weight Heparin in Patients Receiving Chemotherapy for Cancer.***

| Outcome after 12 Months | Participants | Relative Risk (95% CI) | Anticipated Absolute Effect | |
|---|---|---|---|---|
| | | | Risk without LMWH | Risk Difference with LMWH (95% CI) |
| | *no. (no. of studies)* | | *no. of events per 1000 patients* | |
| Death | 6245 (10) | 0.94 (0.88–1.00) | 501 | 30 fewer (60 fewer to 0 more) |
| Symptomatic VTE | 5979 (9) | 0.57 (0.40–0.81) | 46 | 20 fewer (27 fewer to 9 fewer) |
| Major bleeding | 6518 (11) | 1.06 (0.71–1.57) | 16 | 1 more (5 fewer to 9 more) |
| Minor bleeding | 6020 (9) | 1.18 (0.89–1.55) | 27 | 5 more (3 fewer to 15 more) |

Akl & Schünemann, New Engl J Med, 2012

# Do you have confidence in these estimates of effects?

**Table 1. Summary of Findings Table Showing the Relative Risks and Absolute Effects over 12 Months for Each Important Outcome after Treatment with a Low-Molecular-Weight Heparin in Patients Receiving Chemotherapy for Cancer.***

| Outcome after 12 Months | Participants | Relative Risk (95% CI) | Anticipated Absolute Effect | |
|---|---|---|---|---|
| | | | Risk without LMWH | Risk Difference with LMWH (95% CI) |
| | *no. (no. of studies)* | | *no. of events per 1000 patients* | |
| Death | 6245 (10) | 0.94 (0.88–1.00) | 501 | 30 fewer (60 fewer to 0 more) |
| Symptomatic VTE | 5979 (9) | 0.57 (0.40–0.81) | 46 | 20 fewer (27 fewer to 9 fewer) |
| Major bleeding | 6518 (11) | 1.06 (0.71–1.57) | 16 | 1 more (5 fewer to 9 more) |
| Minor bleeding | 6020 (9) | 1.18 (0.89–1.55) | 27 | 5 more (3 fewer to 15 more) |


← 100% confident

← 0% confident

Akl & Schünemann, New Engl J Med, 2012

# High confidence in the effects

**Table 1. Summary of Findings Table Showing the Relative Risks and Absolute Effects over 12 Months for Each Important Outcome after Treatment with a Low-Molecular-Weight Heparin in Patients Receiving Chemotherapy for Cancer.***

| Outcome after 12 Months | Participants | Relative Risk (95% CI) | Anticipated Absolute Effect | | Quality of Evidence (GRADE) and Comments† |
|---|---|---|---|---|---|
| | | | Risk without LMWH | Risk Difference with LMWH (95% CI) | |
| | *no. (no. of studies)* | | *no. of events per 1000 patients* | | |
| Death | 6245 (10) | 0.94 (0.88–1.00) | 501 | 30 fewer (60 fewer to 0 more) | Moderate-quality evidence owing to imprecision and concern about publication bias; a survival analysis based on data from 9 studies shows a hazard ratio of 0.83 (95% CI, 0.72–0.95) |
| Symptomatic VTE | 5979 (9) | 0.57 (0.40–0.81) | 46 | 20 fewer (27 fewer to 9 fewer) | High-quality evidence; the data are combined for pulmonary embolism and symptomatic deep venous thrombosis |
| Major bleeding | 6518 (11) | 1.06 (0.71–1.57) | 16 | 1 more (5 fewer to 9 more) | Moderate-quality evidence owing to imprecision; the increase may be acceptable to patients, given that VTE, which occurs more frequently, may be equally unpleasant |
| Minor bleeding | 6020 (9) | 1.18 (0.89–1.55) | 27 | 5 more (3 fewer to 15 more) | Moderate-quality evidence owing to imprecision; however, this outcome is unlikely to be critical for decision making |

Akl & Schünemann, New Engl J Med, 2012

# A systematic review of the literature on the treatment of pityriasis rubra pilaris type 1 with TNF-antagonists

G. Petrof,* N. Almaani, C.B. Archer

St John's Institute of Dermatology,

*Correspondence: G. Petrof. E-

How confident are you in the estimates of effect?


← 100% confident

← 0% confident

## Abstract

**Background** Adult pityriasi
and histological parallels with pso
tumour necrosis factor (TNF) antagonists.

**Objectives** Our objective was to systematically review the literature for eviden
the treatment of adult PRP.

**Methods** We performed a systematic search of the Cochrane library, EMBASE
We defined diagnosis of PRP, classified clinical response and whether th
antagonists. We also reviewed disease, treatment duration and follow up.

**Results** Sixteen articles were selected for detailed review. From these, 12 art
criteria and were included in the systematic review. The authors identified tv
archive. A total of 15 evaluable cases were included for analysis. Twelve showe
TNF-antagonists with a mean time to maximal response of 5 months. In 10 of t
attributable to TNF antagonist therapy.

**Conclusion** These data indicate that TNF-antagonists may be of value in treating adult type 1 PRP refractory to other systemic agents but selective reporting bias, together with the lack of standard diagnostic criteria and established spontaneous resolution in PRP, prevent any firm recommendations on their place in management.

Received: 10 November 2011; Accepted: 13 January 2012

*JEADV* 2012

# Confidence in estimates of effect

100% confident →


0% confident →

← starting point?

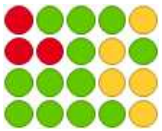
**Bradford Hill Criteria**

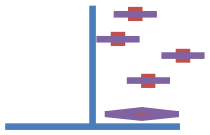
Strength
Consistency
Temporality
Biological gradient
Specificity
Biological Plausibility
Coherence
Experiment
Analogy

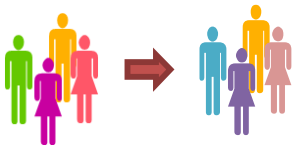
**Good but insufficient (publication bias?)**

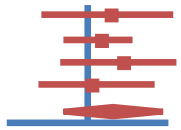
# Determinants of confidence

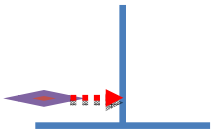
- **RCTs** ⊕⊕⊕⊕
- **observational studies** ⊕⊕◯◯
- **5 factors that can lower quality**
  1. limitations in detailed study design and execution *(risk of bias criteria)*
  2. Inconsistency *(or heterogeneity)*
  3. Indirectness *(PICO and applicability)*
  4. Imprecision
  5. Publication bias
- **3 factors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

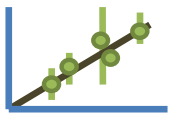












McMaster University
Inspiring Innovation and Discovery

# GRADE Quality of Evidence

In the context of making recommendations:

"The quality of evidence reflects the extent of our confidence that the estimates of an effect are adequate to support a particular decision or recommendation."


***Figure 1.*** *Belief and confidence: a two-dimensional weather report. (Reprinted by permission from the Wall Street Journal).*

Likelihood of and confidence in an outcome


Likelihood of and confidence in an effect

I figure there is a 20% reduction in risk with this intervention and low certainty we know what we are talking about

# GRADE Working Group

McMaster University
Inspiring Innovation and Discovery

## Grades of Recommendation Assessment, Development and Evaluation

- Aim: to develop a common, transparent and sensible system for grading the quality of evidence and the strength of recommendations (over 100 systems)
- International group of guideline developers, epidemiologists, clinical researchers, public health officers, methodologists & clinicians from around the world (>300 contributors) – since 2000
- Over 70 major organizations adopted GRADE

WHO IDSA ATS SIGN Scottish Intercollegiate Guidelines Network THE COCHRANE COLLABORATION NHS National Institute for Health and Clinical Excellence

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

# Assessing Quality of Evidence by Outcome

McMaster University
Inspiring Innovation and Discovery

**Table: GRADE's approach to rating quality of evidence (aka confidence in effect estimates)**

*For each outcome based on a systematic review and across outcomes (lowest quality across the outcomes critical for decision making)*

| 1. Establish initial level of confidence | | 2. Consider lowering or raising level of confidence | | 3. Final level of confidence rating |
|---|---|---|---|---|
| ***Study design*** | ***Initial confidence in an estimate of effect*** | ***Reasons for considering lowering or raising confidence*** | | ***Confidence in an estimate of effect across those considerations*** |
| | | **⬇ Lower if** | **⬆ Higher if*** | |
| ***Randomized trials*** ➔ | **High confidence** | **Risk of Bias**<br>**Inconsistency**<br>**Indirectness**<br>**Imprecision**<br>**Publication bias** | **Large effect**<br>**Dose response**<br>**All plausible confounding & bias**<br>• would reduce a demonstrated effect<br>**or**<br>• would suggest a spurious effect if no effect was observed | **High** ⨁⨁⨁⨁ |
| | | | | **Moderate** ⨁⨁⨁◯ |
| ***Observational studies*** ➔ | **Low confidence** | | | **Low** ⨁⨁◯◯ |
| | | | | **Very low** ⨁◯◯◯ |

*upgrading criteria are usually applicable to observational studies only.

# Lowering confidence in RCTs

**Table: GRADE's approach to rating quality of evidence (aka confidence in effect estimates)**

*For each outcome based on a systematic review and across outcomes (lowest quality across the outcomes critical for decision making)*

| 1. Establish initial level of confidence | | 2. Consider lowering or raising level of confidence | | 3. Final level of confidence rating |
|---|---|---|---|---|
| *Study design* | *Initial confidence in an estimate of effect* | *Reasons for considering lowering or raising confidence* | | *Confidence in an estimate of effect across those considerations* |
| | | ↓ Lower if | ↑ Higher if* | |
| *Randomized trials* → | High confidence | Risk of Bias<br>Inconsistency<br>Indirectness<br>Imprecision<br>Publication bias | Large effect<br>Dose response<br>All plausible confounding & bias<br>• would reduce a demonstrated effect or<br>• would suggest a spurious effect if no effect was observed | High<br>⊕⊕⊕⊕ |
| | | | | Moderate<br>⊕⊕⊕◯ |
| *Observational studies* → | Low confidence | | | Low<br>⊕⊕◯◯ |
| | | | | Very low<br>⊕◯◯◯ |

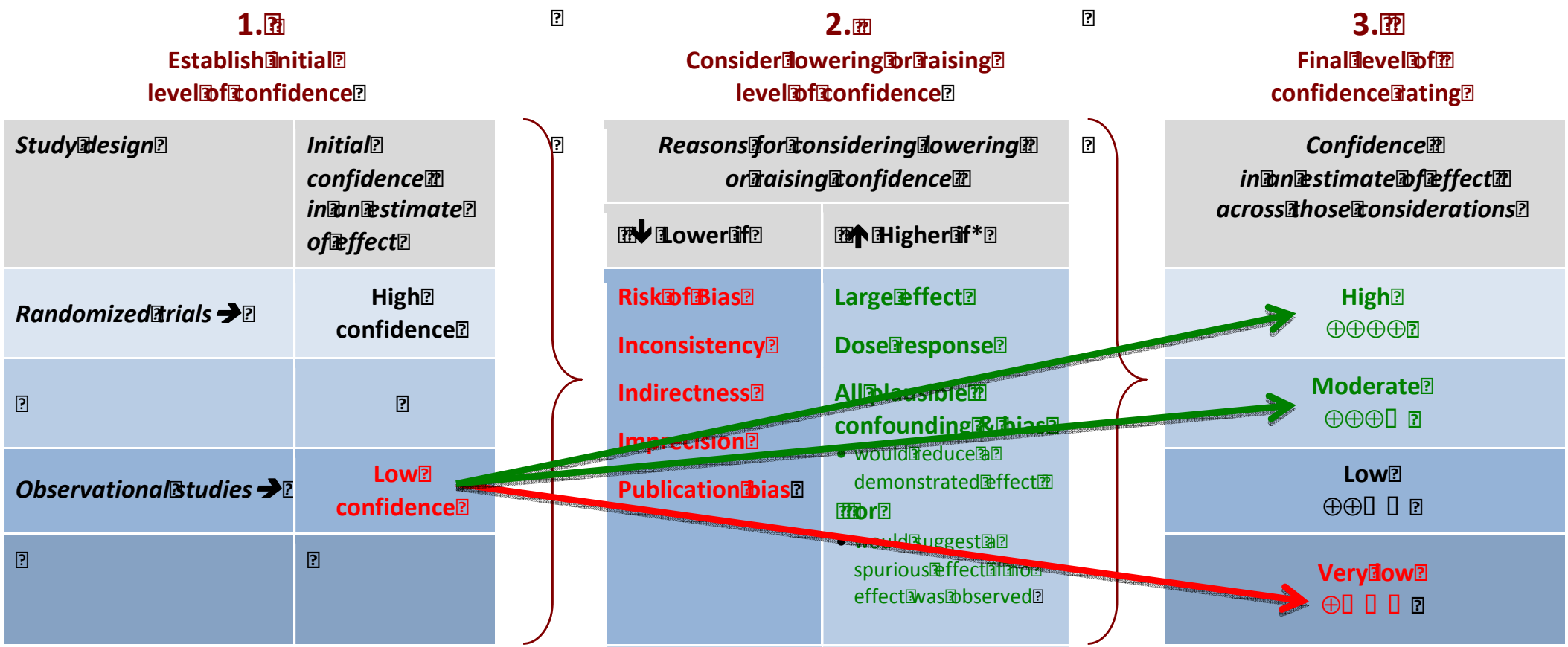
*upgrading criteria are usually applicable to observational studies only.

# Altering confidence in observational studies

McMaster University
Inspiring Innovation and Discovery

**Table: GRADE's approach to rating quality of evidence (aka confidence in effect estimates)**

*For each outcome based on a systematic review and across outcomes (lowest quality across the outcomes critical for decision making)*




*upgrading criteria are usually applicable to observational studies only.

# High confidence in the effects

**Table 1. Summary of Findings Table Showing the Relative Risks and Absolute Effects over 12 Months for Each Important Outcome after Treatment with a Low-Molecular-Weight Heparin in Patients Receiving Chemotherapy for Cancer.***

| Outcome after 12 Months | Participants | Relative Risk (95% CI) | Anticipated Absolute Effect | | Quality of Evidence (GRADE) and Comments† |
|---|---|---|---|---|---|
| | | | Risk without LMWH | Risk Difference with LMWH (95% CI) | |
| | *no. (no. of studies)* | | *no. of events per 1000 patients* | | |
| | | | | | venous thrombosis |
| Major bleeding | 6518 (11) | 1.06 (0.71–1.57) | 16 | 1 more (5 fewer to 9 more) | Moderate-quality evidence owing to imprecision; the increase may be acceptable to patients, given that VTE, which occurs more frequently, may be equally unpleasant |
| Minor bleeding | 6020 (9) | 1.18 (0.89–1.55) | 27 | 5 more (3 fewer to 15 more) | Moderate-quality evidence owing to imprecision; however, this outcome is unlikely to be critical for decision making |

# Should every cancer patient receive heparin?

Akl & Schünemann, New Engl J Med, 2012

# Factors influencing recommendation or decision

- How large is the actual benefit?
- Values and preferences
- Cost or resource use
  - Opportunity cost
- Availability
- Feasibility
- ….

**Formulate question**

Select outcomes

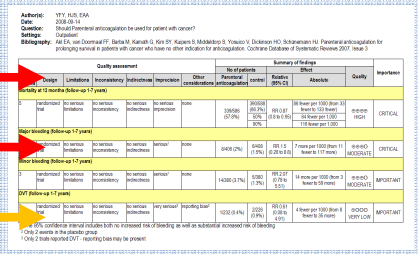
Rate importance

Outcomes across studies

Create evidence profile with GRADEpro

Rate quality of evidence for each outcome

| | | |
|---|---|---|
| P | Outcome | Critical |
| I | Outcome | Critical |
| C | Outcome | Important |
| O | 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

Randomization increases initial quality

Grade up

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

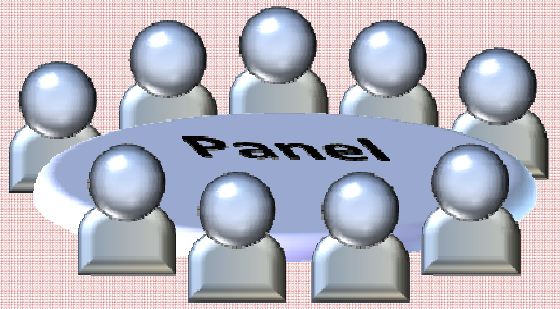
*Systematic review*

Input?

*Recommendation or health care action*

**Grade recommendations**

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




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

By considering balance of:

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

Revise if necessary by considering:

❑Resource use (cost)

**Guideline**


AMERICAN GASTROENTEROLOGICAL ASSOCIATION

Colorectal Cancer Screening and Surveillance: Clinical Guidelines and Rationale—Update Based on New Evidence

**Formulate Recommendations** (↓↑ | ⊕...)

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

**Formulate question**

Select outcomes

Rate importance

Outcomes across studies

Create evidence profile with GRADEpro

Rate quality of evidence for each outcome

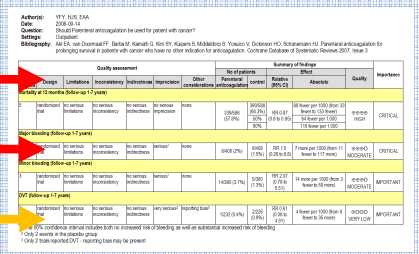
P
I
C
O

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

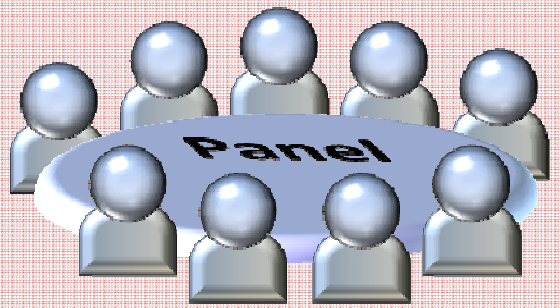
*Systematic review*

*Input?*

*Recommendation or health care action*

**Grade decision**

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




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

By considering balance of:

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

❑Resource use (cost)

**Decision**

**Formulate Recommendations** (↓↑ | ⊕...)

- Cover or not cover
- Research?

McMaster University

Inspiring Innovation and Discovery

# THANK YOU