Cochrane-GRADE Workshop

Modena, June 2017

Holger Schünemann, Elena Parmelli, Sara Balduzzi
Jane Noyes, Heather Munthe-Kaas, Claire Glenton
Background / history of GRADE and GRADE CERQual
Deep vein thrombosis and pulmonary embolism (VTE)

- The risk of VTE is elevated in cancer (4 – 5% annually)
- Require hospital admission and interventions at end of life
- Associated with impairments in function, pain and increased costs
A clinically sensible question

Population: In patients with (lung) cancer, what is the impact of...

Intervention: (comparison) heparin compared with no heparin

Outcomes: on the risk for venous thromboembolism, death, bleeding, burden...?

PICO
A systematic review of RCTs: heparins 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.*

<table>
<thead>
<tr>
<th>Outcome after 12 Months</th>
<th>Participants</th>
<th>Relative Risk (95% CI)</th>
<th>Anticipated Absolute Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>no. (no. of studies)</td>
<td>Risk without LMWH</td>
<td>Risk Difference with LMWH (95% CI)</td>
</tr>
<tr>
<td>Death</td>
<td>6245 (10)</td>
<td>0.94 (0.88–1.00)</td>
<td>501</td>
</tr>
<tr>
<td>Symptomatic VTE</td>
<td>5979 (9)</td>
<td>0.57 (0.40–0.81)</td>
<td>46</td>
</tr>
<tr>
<td>Major bleeding</td>
<td>6518 (11)</td>
<td>1.06 (0.71–1.57)</td>
<td>16</td>
</tr>
<tr>
<td>Minor bleeding</td>
<td>6020 (9)</td>
<td>1.18 (0.89–1.55)</td>
<td>27</td>
</tr>
</tbody>
</table>

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.*

<table>
<thead>
<tr>
<th>Outcome after 12 Months</th>
<th>Participants</th>
<th>Relative Risk (95% CI)</th>
<th>Anticipated Absolute Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>no. (no. of studies)</td>
<td>Risk without LMWH</td>
<td>Risk Difference with LMWH (95% CI)</td>
</tr>
<tr>
<td>Death</td>
<td>6245 (10)</td>
<td>0.94 (0.88–1.00)</td>
<td>501</td>
</tr>
<tr>
<td>Symptomatic VTE</td>
<td>5979 (9)</td>
<td>0.57 (0.40–0.81)</td>
<td>46</td>
</tr>
<tr>
<td>Major bleeding</td>
<td>6518 (11)</td>
<td>1.06 (0.71–1.57)</td>
<td>16</td>
</tr>
<tr>
<td>Minor bleeding</td>
<td>6020 (9)</td>
<td>1.18 (0.89–1.55)</td>
<td>27</td>
</tr>
</tbody>
</table>

Certainty of the evidence?
How confident in the research? GRADE

• Are the research studies well done?
• Are the results consistent across studies when they should be?
• How directly do the results relate to our question?
• Is this effect size precise or likely due to random error?
• Are these all of the studies that have been conducted?
• Plus factors that increase certainty – e.g. large intervention effects
I figure there's a 40% chance of showers and a 10% chance we know what we are talking about.
Systematic review process

1. define the question
2. plan eligibility criteria
3. plan methods
4. search for studies
5. apply eligibility criteria
6. collect data
7. assess studies for risk of bias
8. analyze and present results
9. **interpret results and draw conclusions**
10. improve and update review

Historically not a lot of guidance for this
Cochrane reviews....

...interpret results and draw conclusions?
GRADE criteria *(MECIR standards: mandatory)*
Clinical Practice guidelines & the origin of evidence appraisal system

Effectiveness of intervention

The effectiveness of intervention was graded according to the quality of the evidence obtained, as follows:

I: Evidence obtained from at least one properly randomized controlled trial.

II-1: Evidence obtained from well designed cohort or case-control analytic studies, preferably from more than one centre or research group.

II-2: Evidence obtained from comparisons between times or places with or without the intervention. Dramatic results in uncontrolled experiments (such as the results of the introduction of penicillin in the 1940s) could also be regarded as this type of evidence.

III: Opinions of respected authorities, based on clinical experience, descriptive studies or reports of expert committees.

Classification of recommendations

On the basis of these considerations the task force made a clear recommendation for each condition as to whether it should be specifically considered in a periodic health examination. Recommendations were classified as follows:

A: There is good evidence to support the recommendation that the condition be specifically considered in a periodic health examination.

B: There is fair evidence to support the recommendation that the condition be specifically considered in a periodic health examination.

C: There is poor evidence regarding the inclusion of the condition in a periodic health examination, and recommendations may be made on other grounds.

D: There is fair evidence to support the recommendation that the condition be excluded from consideration in a periodic health examination.

E: There is good evidence to support the recommendation that the condition be excluded from consideration in a periodic health examination.
Rules of Evidence and Clinical Recommendations on the Use of Antithrombotic Agents

D. L. Sackett M.D.

INTRODUCTION

What rules of evidence ought to apply when expert committees meet to generate recommendations for the clinical management of patients? Should only the thoroughly validated results of randomized clinical trials be admissible to avoid or minimize the application of useless or harmful therapy? Or, to maximize the potential benefits to patients (including those possible from unproved remedies), ought a synthesis of the experiences of seasoned clinicians form the basis for such recommendations?

Ample precedent exists for the latter approach even when attempts are made to replace it. However, for the following three reasons, the nonexperimental evidence that forms the recalled experiences of seasoned clinicians will tend to overestimate efficacy:

1. Favorable treatment responses are more likely to be recognized and remembered by clinicians when their patients comply with treatments and keep follow-up appointments. However, there are already five documented instances in which compliant patients in the placebo groups of randomized trials exhibited far more favorable outcomes (including survival) than their noncompliant companions.

Because high compliance is therefore a marker for better outcomes, even when treatment is useless, our uncontrolled clinical experiences often will cause us to conclude that compliant patients must have been receiving efficacious therapy.

2. The natural history of various forms of transient ischemic attacks agents in an effort to halt the progression and complications of thromboembolism. For many of the disorders under consideration here, randomized control trials have never been (and, arguably, never could be) carried out, and the only information base for generating some of the recommendations comes from uncontrolled clinical observations.

What this does mean, however, is that it is important, whenever possible, to base firm recommendations (and especially those involving risk to patients) on the results of rigorously controlled investigations and to be much more circumspect when recommendations rest only on the results of uncontrolled clinical observations. This approach was adopted by the conference participants and led to the definition and adoption of both Levels of Evidence and Grades of Recommendations.

LEVELS OF EVIDENCE

The participants in this undertaking, when summarizing what was known about the causes, clinical course, and management of a given clinical entity, specified the level of evidence that was being used in each case, according to the following classification:

Level 1: Randomized trials with low false-positive (α) and low false-negative (β) errors (high power)

By "low false-positive (α) error" is meant a "positive" trial that demonstrated a statistically significant benefit from experimental treatment. For example, there have now been two randomized trials in which aspirin produced very large, statistically significant reductions in the risk of stroke and death among patients with transient ischemic attacks.

By "low false-negative (β) error (high power)" is meant a "negative" trial that demonstrated no effect of therapy, yet was large enough to exclude the possibility of a clinically important benefit (ie, had very narrow 95% confidence limits that excluded any clinically important improvement from the

Chest 1986
Clinical Recommendations Using Levels of Evidence for Antithrombotic Agents

Deborah J. Cook, MD, FCCP; Gordon H. Guyatt, MD, Chair; Andreas Laupacis, MD; David L. Sackett, MD; and Robert J. Goldberg, PhD

Expert clinical recommendations on the use of antithrombotic agents should be based on the best available evidence. Ideally, this evidence will come from the results of high-quality systematic reviews of rigorously controlled randomized trials.1 Weaker evidence comes from observational studies or uncontrolled clinical experience. Timely implementation of recommendations based on strong evidence can save lives.2 Clinical practice based on the best available literature and recommendations derived from such literature form the foundation of an approach to health care often referred to as evidence-based medicine.3

Chest 1995

<table>
<thead>
<tr>
<th>Level of Evidence</th>
<th>Grade of Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level I</td>
<td>Grade A</td>
</tr>
<tr>
<td>Level I+</td>
<td>Results come from a single RCT in which the lower limit of the CI for the treatment effect exceeds the minimal clinically important benefit</td>
</tr>
<tr>
<td>Level I-</td>
<td>Results come from a meta-analysis of RCTs in which the treatment effects from individual studies are consistent, and the lower limit of the CI for the treatment effect exceeds the minimal clinically important benefit</td>
</tr>
<tr>
<td>Level II</td>
<td>Grade B</td>
</tr>
<tr>
<td>Level II+</td>
<td>Results come from a meta-analysis of RCTs in which the treatment effects from individual studies are consistent and the CI for the treatment effect overlaps the minimal clinically important benefit</td>
</tr>
<tr>
<td>Level II-</td>
<td>Results come from a meta-analysis of RCTs in which the treatment effects from individual studies are widely disparate, but the lower limit of the CI for the treatment effect still exceeds the minimal clinically important benefit</td>
</tr>
<tr>
<td>Level III</td>
<td>Grade C</td>
</tr>
<tr>
<td>Level IV</td>
<td>Grade C</td>
</tr>
<tr>
<td>Level V</td>
<td>Grade C</td>
</tr>
</tbody>
</table>

Results come from nonrandomized concurrent cohort studies

Results come from nonrandomized historic cohort studies

Results come from case series
Grades of Recommendation for Antithrombotic Agents

Gordon Guyatt, MD; Holger Schunemann, MD; Deborah Cook, MD, FCCP; Roman Jaeschke, MD; Stephen Pauker, MD; and Heiner Bucher, MD

Abbreviations: ACCP = American College of Chest Physicians; CI = confidence interval; RCT = randomized, controlled trial; tPA = tissue plasminogen activator

(CHEST 2001; 119:3S–7S)

Treatment decisions involve a trade-off between likely benefits on the one hand, and risks and costs on the other. The Consensus Conference on Antithrombotic Therapy of the American College of Chest Physicians (ACCP) has developed guidelines to help clinicians make supporting studies first and the number denoting the clarity of the risk/benefit trade-off second: 1A, 1B, 1C+, 1C, 2A, and so on. In this iteration, we reflect the primacy of the risk/benefit judgment in determining the recommendation and its strength by placing it first: 1A, 1B, 1C+, 1C, 2A, and so on (Table 1).

The remainder of this article describes the basis of the grading system in more detail. We begin by describing how methodologically strong studies can yield stronger or weaker recommendations depending on the trade-off between risk and benefit.

How Methodologic Quality and Risk Benefit Contribute to Grades of Recommendations
### Grade Practice Recommendations

<table>
<thead>
<tr>
<th>Study quality</th>
<th>Diagnosis</th>
<th>Treatment/prevention/screening</th>
<th>Prognosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1—good-quality patient-oriented evidence</td>
<td>Validated clinical decision rule SR/meta-analysis of high-quality studies High-quality diagnostic cohort studies</td>
<td>SR/meta-analysis of RCTs with consistent findings High-quality individual RCTs All-or-none study§</td>
<td>SR/meta-analysis of good-quality cohort studies Prospective cohort study with good follow-up</td>
</tr>
</tbody>
</table>

#### Level of Certainty

**USPSTF**

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High</strong></td>
<td>The available evidence usually includes consistent results from well-designed, well-conducted studies in representative primary care populations. These studies assess the effects of the preventive service on health outcomes. This conclusion is therefore unlikely to be strongly affected by the results of future studies.</td>
</tr>
<tr>
<td><strong>Moderate</strong></td>
<td>• The available evidence is sufficient to determine the effects of the preventive service on health outcomes, but confidence in the estimate is constrained by such factors as: The number, size, or quality of individual studies. • Inconsistency of findings across individual studies. • Limited generalizability of findings to routine primary care practice. • Lack of coherence in the chain of evidence. As more information becomes available, the magnitude or direction of the observed effect could change, and this change may be large enough to alter the conclusion.</td>
</tr>
<tr>
<td><strong>Low</strong></td>
<td>• The available evidence is insufficient to assess effects on health outcomes. Evidence is insufficient because of: The limited number or size of studies. • Important flaws in study design or methods. • Inconsistency of findings across individual studies. • Gaps in the chain of evidence. • Findings not generalizable to routine primary care practice. • Lack of information on important health outcomes. More information may allow estimation of effects on health outcomes.</td>
</tr>
</tbody>
</table>

#### Consistency across studies

- **Consistent**
  - Most studies or
  - If high-quality evidence favors the preventive service

- **Inconsistent**
  - Considerable or
  - If high-quality evidence favors the preventive service

---

*For evidence and recommendations, please visit [www.plasticsurgery.org](http://www.plasticsurgery.org).*
Which hierarchy?

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

<table>
<thead>
<tr>
<th>Evidence</th>
<th>Recommendation</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>• B</td>
<td>Class I</td>
<td>AHA</td>
</tr>
<tr>
<td>• A</td>
<td>1</td>
<td>ACCP</td>
</tr>
<tr>
<td>• IV</td>
<td>C</td>
<td>SIGN</td>
</tr>
</tbody>
</table>
Simple hierarchies are (too) simplistic

**STUDY DESIGN**

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

**BIAS**

Expert Opinion

Schünemann & Bone, 2003
Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials

Gordon C S Smith, Jill P Pell

Parachutes reduce the risk of injury after gravitational challenge, but their effectiveness has not been proved with randomised controlled trials
Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials

Gordon C S Smith, Jill P Pell

Relative risk reduction:
....> 99.9 % (1/100,000)

U.S. Parachute Association reported 821 injuries and 18 deaths out of 2.2 million jumps in 2007
Aim: to develop a common, transparent and sensible system for grading the quality of evidence and the strength of recommendations
Proposed acronyms  Meeting minutes London 2002

GERIC: Grading Evidence and Recommendation International Collaboration (?)

GRASP: Grades of Recommendation ASsessment and Planning group (Andy)

GRADE: Grades of Recommendation Assessement, Development and Evaluation (Working) Group or "GRADE (Working) Group" in short. (Holger)

One could also use DEsign instead of Development
GRADE: Grades of Recommendation Assessement and Development Enterprise (Holger)

GEAR: Grades of Evidence And Recommendations (Andy)

**Correspondence to:** Dr. Holger J. Schünenemann, Departments of Medicine and of Social and Preventive Medicine, University of Buffalo, 270 Farber Hall, 3435 Main St., Buffalo NY 14214, USA; fax 716-898-4493; hjs@buffalo.edu

**Members of the GRADE Working Group:** David Atkins, Chief Medical Officer, Center for Practice and Technology Assessment, Agency for Healthcare Research and Quality, USA; Dana Best, Assistant Professor, Department of General Pediatrics and Adolescent Medicine, George Washington University, Children’s National Medical Center, USA; Peter A Briss, Acting Chief Community Guide Branch, Centers for Disease Control and Prevention, USA; Martin Eccles, Professor, and James Mason, Professor, Centre for Health Services Research, University of Newcastle upon Tyne, U.K.; Yngve Falck-Ytter, Associate Director, German Cochrane Centre, Institute for Medical Biometry and Medical Informatics, University Hospital Freiburg, Germany; Gunn E. Vist, Researcher, Signe Flottorp, Researcher, and Andrew D. Oxman, Director, Department of Health Services Research, Norwegian Directorate for Health and Social Welfare, Norway; Gordon H. Guyatt, Professor, and Roman Jaeschke, Associate Clinical Professor, Departments of Clinical Epidemiology and Biostatistics and Medicine, McMaster University, Canada; Robin T. Harbour, Quality and Information Director, Scottish Intercollegiate Guidelines Network, United Kingdom; Margaret C. Haugh, Methodologist, Fédération Nationale des Centres de Lutte Contre le Cancer, France; David Henry, Professor and Suzanne Hill, Senior Lecturer, Department of Clinical Pharmacology, Faculty of Medicine and Health Sciences, University of Newcastle, Australia; Gillian Leng, Guidelines Programme Director, National Institute for Clinical Excellence, United Kingdom; Alessandro Liberati, Professor, Università di Modena e Reggio Emilia and Centro per la Valutazione della Efficacia della Assistenza Sanitaria, Italy; Nicola Magrini, Director, Centro per la Valutazione della Efficacia della Assistenza Sanitaria, Italy; Philippa Middleton, Honorary Research Fellow, Australasian Cochrane Centre, Australia; Jacek Mrukowicz, Executive Director, Polish Institute for Evidence Based Medicine, Poland; Dianne O’Connell, Senior Epidemiologist, Cancer Epidemiology Research Unit, Cancer Research and Registers Division, The Cancer Council, Australia; Bob Phillips, Associate Fellow, Centre for Evidence-based Medicine, University Department of Psychiatry, Wameford Hospital, United Kingdom; Holger J Schünenemann, Assistant Professor, Departments of Medicine and of Social & Preventive Medicine, University of Buffalo, USA; Tessa Tan-Torres Edejer, Medical Officer/Scientist, Global Programme on Evidence for Health Policy, World Health Organisation, Switzerland; Helena Varonen, Associate Editor, Finnish Medical Society Duodecim, Finland; John W. Williams Jr., Associate Professor, The Center for Health Services Research in Primary Care, Health Services Research and Development, Department of Veterans Affairs Medical Center and Duke University Medical Center, USA; Stephanie Zaza, Acting Associate Director for Science, Epidemiology Program Office, Centers for Disease Control and Prevention, USA.
Developed a unifying, transparent and sensible system for grading the certainty of evidence and developing recommendations/making decisions

- NICE, WHO, CDC, AHRQ, professional societies, academic institutions
- For systematic reviews, HTA and guidelines
- International contributors (>500) with diversity in background 2008 BMJ series; 2011 JCE series – over 30,000 cites
- Various other publications (incl. GRADE Handbook)
- IT applications GRADEpro GDT

Over 100 organizations adopted or use GRADE
Open membership – free: www.gradeworkinggroup.org
Certainty of the evidence?
How confident in the research? GRADE

• Are the research studies well done?
• Are the results consistent across studies when they should be?
• How directly do the results relate to our question?
• Is this effect size precise or likely due to random error?
• Are these all of the studies that have been conducted?
• Plus factors that increase certainty – e.g. large intervention effects
Determinants of certainty of evidence

- RCTs ★★★★★ | high
- Observational studies ★★★★ | low

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
Assessing Certainty in Evidence by Outcome

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

1. Establish initial level of certainty

<table>
<thead>
<tr>
<th>Study design</th>
<th>Initial certainty of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized trials ➔</td>
<td>High certainty</td>
</tr>
<tr>
<td>Observational studies ➔</td>
<td>Low certainty</td>
</tr>
</tbody>
</table>

2. Consider lowering or raising level of certainty

<table>
<thead>
<tr>
<th>Reasons for considering lowering or raising certainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>↓ Lower if</td>
</tr>
<tr>
<td>Risk of Bias</td>
</tr>
<tr>
<td>Inconsistency</td>
</tr>
<tr>
<td>Indirectness</td>
</tr>
<tr>
<td>Imprecision</td>
</tr>
<tr>
<td>Publication bias</td>
</tr>
<tr>
<td>↑ Higher if*</td>
</tr>
<tr>
<td>Large effect</td>
</tr>
<tr>
<td>Dose response</td>
</tr>
<tr>
<td>All plausible confounding &amp; bias</td>
</tr>
<tr>
<td>• would reduce a demonstrated effect</td>
</tr>
<tr>
<td>• would suggest a spurious effect if no effect was observed</td>
</tr>
</tbody>
</table>

3. Final level of certainty rating

<table>
<thead>
<tr>
<th>Certainty of the evidence across those considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
</tr>
<tr>
<td>Moderate</td>
</tr>
<tr>
<td>Low</td>
</tr>
<tr>
<td>Very low</td>
</tr>
</tbody>
</table>

*upgrading criteria are usually applicable to observational studies only.
Lowering certainty in RCTs

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

<table>
<thead>
<tr>
<th>Study design</th>
<th>Initial certainty of evidence</th>
<th>Reasons for considering lowering or raising level of certainty</th>
<th>Certainty of the evidence across those considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized trials</td>
<td>High certainty</td>
<td>Lower if: Risk of Bias, Inconsistency, Indirectness, Imprecision, Publication bias</td>
<td>High ⬆️</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Higher if*: Large effect, Dose response</td>
<td>Moderate ⬇️</td>
</tr>
</tbody>
</table>
| Observational studies        | Low certainty                 | *All plausible confounding & bias:*
|                              |                               | • would reduce a demonstrated effect or
|                              |                               | • would suggest a spurious effect if no effect was observed | Low ⬇️ |
|                              |                               | *upgrading criteria are usually applicable to observational studies only.* | Very low ⬇️ |

1. Establish initial level of certainty
2. Consider lowering or raising level of certainty
3. Final level of certainty rating
Altering certainty of non-randomized studies

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

1. Establish initial level of certainty

<table>
<thead>
<tr>
<th>Study design</th>
<th>Initial certainty of the evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized trials</td>
<td>High confidence</td>
</tr>
<tr>
<td>Observational studies</td>
<td>Low confidence</td>
</tr>
</tbody>
</table>

2. Consider lowering or raising level of certainty

<table>
<thead>
<tr>
<th>Reasons for considering lowering or raising certainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>↓ Lower if</td>
</tr>
<tr>
<td>Risk of Bias</td>
</tr>
<tr>
<td>Inconsistency</td>
</tr>
<tr>
<td>Indirectness</td>
</tr>
<tr>
<td>Imprecision</td>
</tr>
<tr>
<td>Publication bias</td>
</tr>
<tr>
<td>↑ Higher if*</td>
</tr>
<tr>
<td>Large effect</td>
</tr>
<tr>
<td>Dose response</td>
</tr>
<tr>
<td>All plausible confounding &amp; bias</td>
</tr>
<tr>
<td>- would suggest a spurious effect if no effect was observed</td>
</tr>
</tbody>
</table>

3. Final level of certainty rating

<table>
<thead>
<tr>
<th>Certainty of the evidence across those considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
</tr>
<tr>
<td>Moderate</td>
</tr>
<tr>
<td>Low</td>
</tr>
<tr>
<td>Very low</td>
</tr>
</tbody>
</table>

*upgrading criteria are usually applicable to observational studies only.
• **Certainty of evidence**
  – Assess evidence transparently across all domains
  – Confidence in an estimate?
  – Starts with single research studies
  – Ends with a body of evidence by health outcome
    • high, moderate, low, very low certainty

• **Recommendations/Decisions**
  – Involves making judgments and decisions transparent
  – Evidence to Decision (EtD) frameworks
    • Comprehensive list of criteria that influence a decision or recommendation
  – Clearly developed & formulated action message
    • Strong or conditional recommendations for or against an option
Recommendation/Decision

Grade recommendations (Evidence to Recommendation)
• For or against (direction) ↓↑
• Strong or conditional/weak (strength)

By considering balance of consequences (evidence to recommendations):
- Quality of evidence
- Balance benefits/harms
- Values and preferences (equity)
- Resource use (cost, feasibility)
- Acceptability

Guideline/Decision

Formulate Recommendations (↓↑ | ⊕…)
“The panel recommends that …should…”
“The panel suggests that …should…”
“The panel suggests to not …”
“The panel recommends to not…”

Transparency, clear, actionable

EtD framework

Evidence synthesis (systematic review/HTA)

Create evidence profile/SoF Table with GRADEpro

Summary of findings & estimate of effect for each outcome

Rate quality of evidence for each outcome
Randomization raises initial quality
RCTs: high
Observational: low

Grade overall quality of evidence across lowest quality of critical outcomes

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

Outcomes across studies

Formulate question
Select outcomes
Rate importance

P I/E C O

Outcome Critical
Outcome Critical
Outcome Important
Outcome Not important

Input

Critical Important

Very low
Low
Moderate
High

Quality of evidence
Balance benefits/harms
Values and preferences (equity)
Resource use (cost, feasibility)
Acceptability

Panel

GRADE working group

Grade recommendations
(Evidence to Recommendation)
A clinically sensible question

Population: In patients with (lung) cancer, what is the impact of

Intervention: (comparison) heparin compared with no heparin

Outcomes: on the risk for venous thromboembolism, death, bleeding, burden...?

PICO
Anticoagulation for the long-term treatment of venous thromboembolism in patients with cancer
Elie A Akl, Lara A Kahale, Maddalena Barba, Ignacio Neumann, Nawman Labedi, Irene Terrenato, Francesca Sperati, Paola Muti and Holger Schünemann
Online Publication Date: July 2014

Parenteral anticoagulation in ambulatory patients with cancer
Elie A Akl, Lara A Kahale, Rami A Ballout, Maddalena Barba, Victor E D Yosuico, Frederiek F van Doormaal, Saskia Middeldorp, Andrew Bryant and Holger Schünemann
Online Publication Date: December 2014

Low molecular weight heparin versus unfractionated heparin for perioperative thromboprophylaxis in patients with cancer
Elie A Akl, Lara A Kahale, Francesca Sperati, Ignacio Neumann, Nawman Labedi, Irene Terrenato, Maddalena Barba, Elena V Sempos, Paola Muti, Deborah Cook and Holger Schünemann
Online Publication Date: June 2014

Oral anticoagulation in patients with cancer who have no therapeutic or prophylactic indication for anticoagulation
Elie A Akl, Lara Kahale, Irene Terrenato, Ignacio Neumann, Victor E D Yosuico, Maddalena Barba, Francesca Sperati and Holger Schünemann
Online Publication Date: July 2014

Anticoagulation for the initial treatment of venous thromboembolism in patients with cancer
Elie A Akl, Lara A Kahale, Ignacio Neumann, Maddalena Barba, Francesca Sperati, Irene Terrenato, Paola Muti and Holger Schünemann
Online Publication Date: June 2014

Anticoagulation for people with cancer and central venous catheters
Elie A Akl, Elie P Ramly, Lara A Kahale, Victor E D Yosuico, Maddalena Barba, Francesca Sperati, Deborah Cook and Holger Schünemann
Online Publication Date: October 2014
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.*

<table>
<thead>
<tr>
<th>Outcome after 12 Months</th>
<th>Participants</th>
<th>Relative Risk (95% CI)</th>
<th>Anticipated Absolute Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>no. (no. of studies)</td>
<td>Risk without LMWH</td>
<td>Risk Difference with LMWH (95% CI)</td>
</tr>
<tr>
<td>Death</td>
<td>6245 (10)</td>
<td>0.94 (0.88–1.00)</td>
<td>501, 30 fewer (60 fewer to 0 more)</td>
</tr>
<tr>
<td>Symptomatic VTE</td>
<td>5979 (9)</td>
<td>0.57 (0.40–0.81)</td>
<td>46, 20 fewer (27 fewer to 9 fewer)</td>
</tr>
<tr>
<td>Major bleeding</td>
<td>6518 (11)</td>
<td>1.06 (0.71–1.57)</td>
<td>16, 1 more (5 fewer to 9 more)</td>
</tr>
<tr>
<td>Minor bleeding</td>
<td>6020 (9)</td>
<td>1.18 (0.89–1.55)</td>
<td>27, 5 more (3 fewer to 15 more)</td>
</tr>
</tbody>
</table>

Determinants of quality/certainty of a body of evidence

• RCTs 🌟🌟🌟🌟
• observational studies (NRS) 🌟🌟🌟

• 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
### Risk of Bias – Within & Across

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Weber 2008</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Sideras 2006</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Perry 2010</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Peizer 2009</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Lebeau 1994</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Klerk 2005</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Kakkar 2004</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Altinbas 2004</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Agnelli 2008</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>