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

GIN-NA webinar, May 30, 2013

GRADE: Introduction and update

Contributions from the GRADE WG
Disclosure

• Co-chair GRADE Working Group
• World Health Organization: various committees
• Work with various other guideline groups using GRADE
  – E.g. ACP, ACCP
• No direct/personal for profit payments = no financial COI
“Birthplace of evidence-based medicine and problem based learning”
GRADE:

• **Involves assessing evidence transparently:**
  – Confidence in an estimate of effect?
  – Starts with single studies
  – Ends with a body of evidence by outcome and a recommendation

• **Developing structured health care recommendations:**
  – Evidence to recommendation frameworks
  – Comprehensive list of factors that influence a recommendation
  – Clearly developed and formulated message for action
Today’s talk

• Where does GRADE come from and how does it fit in the guideline development process

• Introduction to GRADE
  – Confidence or certainty in evidence (quality)
    • Examples
  – Evidence to recommendations
    • Examples

• Updates & Developments in GRADE
  – New online tool, other resources
  – Diagnosis
  – Prognosis
Literature

American Thoracic Society Documents

A Guide to Guidelines for Professional Societies and Other Developers of Recommendations
Introduction to Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report

Holger J. Schünemann, Mark Woodhead, Antonio Anzueto, A. Sonia Buist, William MacNee, Klaus F. Rabe, and John Heffner; on behalf of the ATS and Coordinating Efforts in COPD Guideline Development Task Force.

Health Research Policy and Systems

Review
Improving the use of research evidence in guideline development: introduction
Andrew D Oxman¹, Atle Fretheim¹, Holger J Schünemann² and SURE³

Journal of Clinical Epidemiology

GRADE SERIES - GUEST EDITORS, SHARON STRAUS AND SASHA SHEPHERD
GRADE guidelines: A new series of articles in the Journal of Clinical Epidemiology
Gordon H. Guyatt¹,²,³, Andrew D. Oxman³, Holger J. Schünemann¹,², Peter Tugwell⁴, Andre Knottnerus⁵
¹Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario L8N 3Z5, Canada
²Department of Medicine, McMaster University, Hamilton, Ontario L8N 3Z5, Canada
³Norwegian Knowledge Centre for the Health Services, PO Box 7004, St Olavs plass, 0130 Oslo, Norway
⁴Centre for Global Health, Institute of Population Health, University of Ottawa, Ottawa, Ontario K1N 6N5, Canada
⁵Department of General Practice, PO Box 616, 6200 MD Maastricht, The Netherlands
Accepted 29 September 2010

Implementation Science

METHODOLOGY
Developing clinical practice guidelines: target audiences, identifying topics for guidelines, guideline group composition and functioning and conflicts of interest
Martin P Eccles, Jeremy M Grimshaw, Paul Shekelle, Holger J Schünemann and Steven Woolf

Elsevier
Resources

- Cebgrade.mcmaster.ca
- www.gradeworkinggroup.org
- YouTube
- GRADEpro tool (including detailed handbook):
  http://ims.cochrane.org/gradeapro
Guideline enterprise

1. Priority Setting
2. Planning
3. Guideline Development Group Composition
4. Establishing Guideline Group Process
5. Identifying Target Audience and Topic Selection
6. Stakeholder Involvement
7. Consumer Involvement
8. Conflict of Interest Considerations
9. Question Generation (PICO incl. outcome selection)
10. Deciding what Evidence to Include & Searching
11. Appraising Quality of Evidence
12. Summarizing Evidence
13. Developing Recommendations
14. Determining Strength of Recommendations
15. Wording of Recommendations
16. Reporting and Peer Review
17. Dissemination and Implementability
18. Updating the Guideline
19. Evaluating the Guideline Development
Evidence synthesis

Formulate question
- Select outcomes
- Rate importance

Outcomes across studies

Create evidence profile with GRADEpro

Rate quality of evidence for each outcome

Summary of findings & estimate of effect for each outcome

Randomization increases initial quality

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

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

Grade down
0
1
2
3
4
5

Grade up
Very low
Low
Moderate
High

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

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)

Recommendation

Evidence synthesis
Implications of a strong recommendation

• Policy makers: The recommendation can be adapted as a policy in most situations

• Patients: Most people in this situation would want the recommended course of action and only a small proportion would not

• Clinicians: Most patients should receive the recommended course of action
Implications of a weak recommendation

- Policy makers: There is a need for substantial debate and involvement of stakeholders
- Patients: The majority of people in this situation would want the recommended course of action, but many would not
- Clinicians: Be more prepared to help patients to make a decision that is consistent with their own values/decision aids and shared decision making
Evidence synthesis

Formulate question
- Select outcomes
- Rate importance

Outcomes across studies
- Create evidence profile with GRADEpro
- Rate quality of evidence for each outcome

Randomization increases initial quality

Grade down
- Risk of bias
- Inconsistency
- Indirectness
- Imprecision
- Publication bias

Grade up
- Large effect
- Dose response
- Opposing bias & Confounders

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

Recommendation

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)

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

Guideline
Separation of judgments:
1) 4 categories of quality of evidence:
   – methodological quality of evidence
   – likelihood of bias related to recommendation
   – by outcome and across outcomes
2) Recommendation:
   • 2 grades
   • Weak (aka conditional) or strong (for or against an action)
     – balance of benefits and downsides
     – values and preferences
     – resource use
     – quality of evidence
From evidence to recommendations

- RCTs
- Observational studies

- High level recommendation
- Lower level recommendation

Old systems

GRADE

- Balance between benefits, harms & burdens
- Quality of evidence
- Resource use

Patients’ values & preferences
Pancreatic cancer

- Ninth most common cancer
- 277,000 new cases annually worldwide
- Etiology?

- Do non-steroidal anti-inflammatory drugs (NSAIDS) increase the risk of pancreatic cancer?
- In patients with arthritis does therapy with NSAIDS increase the risk of pancreatic cancer?

Ferlay J, Shin HR, Bray F, et al
**Systematic review: Non-steroidal drug use and risk of pancreatic cancer**

<table>
<thead>
<tr>
<th>Study or sub-category</th>
<th>ASA/NSAIDs use $n/N$</th>
<th>No/occasional use $n/N$</th>
<th>OR (random) $95%$ CI</th>
<th>OR (random) $95%$ CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (95% CI)</td>
<td>79,761</td>
<td>831,811</td>
<td>1.11 (0.84, 1.47)</td>
<td>1.11 (0.84, 1.47)</td>
</tr>
</tbody>
</table>

Total events: 486 (NSAIDS, 4396 control)

*Meta-analysis: the use of non-steroidal anti-inflammatory drugs and pancreatic cancer risk for different exposure categories.*  
### Systematic review: Non-steroidal drug use and risk of pancreatic cancer

**Capurso G, Schünemann HJ, Terrenato I, et al.**

*Meta-analysis: the use of non-steroidal anti-inflammatory drugs and pancreatic cancer risk for different exposure categories.*  

<table>
<thead>
<tr>
<th>Study or sub-category</th>
<th>ASA/NSAIDs use</th>
<th>No/occasional use</th>
<th>OR (random) 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n/N</td>
<td>n/N</td>
<td></td>
</tr>
<tr>
<td>Anderson</td>
<td>10/6012</td>
<td>60/17 277</td>
<td>0.48 [0.24, 0.93]</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td>79,761</td>
<td>831,811</td>
<td>1.11 (0.84, 1.47)</td>
</tr>
</tbody>
</table>

Total events: 486 (NSAIDS, 4396 control)
### Systematic review: Non-steroidal drug use and risk of pancreatic cancer

#### Meta-analysis: the use of non-steroidal anti-inflammatory drugs and pancreatic cancer risk for different exposure categories.


<table>
<thead>
<tr>
<th>Study or sub-category</th>
<th>ASA/NSAIDs use n/N</th>
<th>No/occasional use n/N</th>
<th>OR (random) 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson</td>
<td>10/6012</td>
<td>60/17 277</td>
<td>0.48 [0.24, 0.93]</td>
</tr>
<tr>
<td>Schernhammer</td>
<td>37/13 284</td>
<td>153/89 541</td>
<td>1.63 [1.14, 2.34]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>79,761</td>
<td>831,811</td>
<td>1.11 (0.84, 1.47)</td>
</tr>
</tbody>
</table>

Total events: 486 (NSAIDS, 4396 control)
Systematic review: Non-steroidal drug use and risk of pancreatic cancer

Meta-analysis: the use of non-steroidal anti-inflammatory drugs and pancreatic cancer risk for different exposure categories.


Total events: 486 (NSAIDS, 4396 control)
Test for heterogeneity: $X^2$ p = 0.001, $I^2$ = 75%
I figure there’s a 40% chance of showers and a 10% chance we know what we are talking about.

Likelihood of and confidence in the effect
I figure there is a 20% reduction in risk with this intervention and a 10% chance we know what we are talking about.
Confidence in estimates of effect

Bradford Hill Criteria
- Strength
- Consistency
- Temporality
- Biological gradient
- Specificity
- Biological Plausibility
- Coherence
- Experiment
- Analogy

Good but insufficient (publication bias?)
The origin of evidence appraisal systems

The periodic health examination

Canadian Task Force on the Periodic Health Examination*

*Editor-in-Chief, CMA J. 1979; Vol. 121, 1193.
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.
Scientific Evidence Underlying the ACC/AHA Clinical Practice Guidelines

Canadian Task Force 1979

Level of evidence C: recommendations based on expert opinion, case studies OR standards of care.

Conclusions

Recommendations issued in current ACC/AHA clinical practice guidelines are largely developed from lower levels of evidence or expert opinion. The proportion of recommendations for which there is no conclusive evidence is also growing. These findings highlight the need to improve the process of writing guidelines and to expand the evidence base from which clinical practice guidelines are derived.

Context

The joint cardiovascular practice guidelines of the American College of Cardiology (ACC) and the American Heart Association (AHA) have become important documents for guiding cardiology practice and establishing benchmarks for quality of care.

Objective

To describe the evolution of recommendations in ACC/AHA cardiovascular guidelines and the distribution of recommendations across classes of recommendations and levels of evidence.

Level of evidence C: recommendations based on expert opinion, case studies OR standards of care.

III: Opinions of respected authorities, based on clinical experience, descriptive studies or reports of expert committees.
<table>
<thead>
<tr>
<th>Guidelines</th>
<th>Year</th>
<th>Class of Recommendations</th>
<th>Level of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>Atrial fibrillation^7</td>
<td>2006</td>
<td>41/111 (36.9)</td>
<td>55/111 (49.5)</td>
</tr>
<tr>
<td>Heart failure^33</td>
<td>2005</td>
<td>66/129 (51.2)</td>
<td>44/129 (41.1)</td>
</tr>
<tr>
<td>Peripheral artery disease^33</td>
<td>2005</td>
<td>147/237 (62.6)</td>
<td>68/237 (28.1)</td>
</tr>
<tr>
<td>STEMI^45</td>
<td>2004</td>
<td>248/422 (58.8)</td>
<td>123/422 (29.1)</td>
</tr>
<tr>
<td>Perioperative evaluation^40</td>
<td>2007</td>
<td>13/50 (26.0)</td>
<td>27/50 (54.0)</td>
</tr>
<tr>
<td>Secondary prevention^44</td>
<td>2006</td>
<td>38/48 (79.2)</td>
<td>10/48 (20.8)</td>
</tr>
<tr>
<td>Stable angina^47</td>
<td>2002</td>
<td>78/235 (33.2)</td>
<td>98/235 (41.7)</td>
</tr>
<tr>
<td>Supraventricular arrhythmias^48</td>
<td>2003</td>
<td>61/147 (41.5)</td>
<td>77/147 (52.4)</td>
</tr>
<tr>
<td>Unstable angina^51</td>
<td>2007</td>
<td>137/298 (46.2)</td>
<td>82/298 (27.5)</td>
</tr>
<tr>
<td>Valvular heart disease^55</td>
<td>2008</td>
<td>156/320 (48.8)</td>
<td>124/320 (38.8)</td>
</tr>
<tr>
<td>Ventricular arrhythmias and sudden cardiac death^52</td>
<td>2006</td>
<td>103/217 (47.5)</td>
<td>100/217 (46.1)</td>
</tr>
<tr>
<td>Summary of disease guidelines, median (IQR), %</td>
<td>48.8</td>
<td>38.8</td>
<td>12.1</td>
</tr>
</tbody>
</table>
What we had learned

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

Parachutes reduce the risk of injury after gravitational challenge, but their effectiveness has not been proved with randomised trials

BMJ, 2003
When are randomised trials unnecessary? Picking signal from noise

The relation between a treatment and its effect is sometimes so dramatic that bias can be ruled out as an explanation. Paul Glasziou and colleagues suggest how to determine when observations speak for themselves.

Some historical examples of treatments with dramatic effects

- Insulin for diabetes
- Blood transfusion for severe haemorrhagic shock
- Sulphanilamide for puerperal sepsis
- Streptomycin for tuberculous meningitis
- Defibrillation for ventricular fibrillation
- Closed reduction and splinting for fracture of long bones with displacement
- Salicin for acute rheumatism
- Neostigmine for myasthenia gravis
- Tracheostomy for tracheal obstruction
- Suturing for repairing large wounds
- Drainage for pain associated with abscesses
- Pressure or suturing for arresting haemorrhage
- Ether for anaesthesia
- One way valve or underwater seal drainage for pneumothorax and haemothorax
- Phototherapy for skin tuberculosis
- Combination chemotherapy with cisplatin, vinblastine, and bleomycin for disseminated testicular cancer
Background: GRADE

- **GRADE working group** - International contributors (>300) with diversity in background – since year 2000
- Developed a unifying, transparent and sensible system for grading the quality of evidence and developing recommendations
- Over 70 organizations adopted or use GRADE
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.”
Confidence in estimates of effect

100% confident $\rightarrow$GRADE’s starting point$\leftarrow$ 0% confident $\rightarrow$
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
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>

Confidence in the estimates

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>

Determinants of confidence of a body of evidence

- **RCTs** ⊕⊕⊕⊕⊕
- **observational studies** ⊕⊕ΟΟΟ

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>
</tr>
</thead>
<tbody>
<tr>
<td>Agnelli 2009</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>
</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>
</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>
</tr>
<tr>
<td>Pelzer 2009</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>
</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>
</tr>
</tbody>
</table>
Imprecision
Inconsistency

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Heparin Events</th>
<th>Heparin Total</th>
<th>Control Events</th>
<th>Control Total</th>
<th>Weight (%)</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altinbas 2004</td>
<td>0</td>
<td>42</td>
<td>1</td>
<td>42</td>
<td>1.2%</td>
<td>0.33 [0.01, 7.96]</td>
</tr>
<tr>
<td>Agnelli 2009</td>
<td>11</td>
<td>769</td>
<td>11</td>
<td>381</td>
<td>13.4%</td>
<td>0.50 [0.22, 1.13]</td>
</tr>
<tr>
<td>Perry 2010</td>
<td>11</td>
<td>99</td>
<td>14</td>
<td>87</td>
<td>15.8%</td>
<td>0.69 [0.33, 1.44]</td>
</tr>
<tr>
<td>Pelzer 2009</td>
<td>8</td>
<td>160</td>
<td>22</td>
<td>152</td>
<td>14.6%</td>
<td>0.35 [0.16, 0.75]</td>
</tr>
<tr>
<td>Weber 2008</td>
<td>1</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>1.3%</td>
<td>3.00 [0.14, 65.90]</td>
</tr>
<tr>
<td>Sideras 2006</td>
<td>4</td>
<td>68</td>
<td>5</td>
<td>70</td>
<td>6.7%</td>
<td>0.82 [0.23, 2.94]</td>
</tr>
<tr>
<td>Agnelli 2012</td>
<td>20</td>
<td>1608</td>
<td>55</td>
<td>1604</td>
<td>24.6%</td>
<td>0.36 [0.22, 0.60]</td>
</tr>
<tr>
<td>van Doormaal 2011</td>
<td>16</td>
<td>244</td>
<td>15</td>
<td>259</td>
<td>17.4%</td>
<td>1.13 [0.57, 2.24]</td>
</tr>
<tr>
<td>Kakkar 2004</td>
<td>3</td>
<td>190</td>
<td>4</td>
<td>184</td>
<td>5.1%</td>
<td>0.73 [0.16, 3.20]</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>3190</strong></td>
<td><strong>2789</strong></td>
<td><strong>100.0%</strong></td>
<td><strong>Risk Ratio</strong></td>
<td><strong>0.57 [0.40, 0.81]</strong></td>
<td></td>
</tr>
</tbody>
</table>

Total events: 74 | 127

Heterogeneity: Tau^2 = 0.07; Chi^2 = 10.50, df = 8 (P = 0.73); I^2 = 24%
Test for overall effect: Z = 3.13 (P = 0.002)

Publication bias
Indirectness
Relation between PICO and available evidence
Relation between PICO and available evidence

Judgment: good representation of patients with advanced, incl. pancreatic cancer
**Table 1. Summary of Treatment with a Low-Molecular Weight Heparin for 12 Months for Each Important Outcome after Treatment for Cancer.**

<table>
<thead>
<tr>
<th>Outcome after 12 Months</th>
<th>Effect</th>
<th>95% CI</th>
<th>No. of Events per 1000 Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Death</td>
<td>0.94</td>
<td>0.88–1.00</td>
<td>501 (60 fewer to 0 more)</td>
</tr>
<tr>
<td>Symptomatic VTE</td>
<td>0.57</td>
<td>0.40–0.81</td>
<td>46 (27 fewer to 9 fewer)</td>
</tr>
<tr>
<td>Major bleeding event</td>
<td>0.64</td>
<td>0.40–0.99</td>
<td>3 (0 fewer to 3 more)</td>
</tr>
</tbody>
</table>

No downgrading:
- Low Risk of bias
- Little inconsistency
- Little imprecision
- Undetected publication bias
- Little indirectness

Quality remains high

**What about mortality?**

<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>
</tr>
</thead>
<tbody>
<tr>
<td>Agnelli 2009</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>
</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>
</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>
</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>
</tr>
<tr>
<td>Pelzer 2009</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>
</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>
</tr>
<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>
</tr>
</tbody>
</table>
1. Limitations in detailed study design and execution (risk of bias criteria)
2. Indirectness (PICO and applicability)
3. Inconsistency (or heterogeneity)
4. Imprecision
5. Publication bias

Explanation?

- \( I^2 \)
- P-value
- Overlap in CI
- Widely differing point estimates
- Small sample size & small number of events
- Wide confidence intervals
- Uncertainty about magnitude of effect

Suggested Rules in handbook
Publication bias
Indirectness: Relation between PICO and available evidence
Non-randomized studies as a source of complementary, sequential or replacement evidence for randomized controlled trials in systematic reviews on the effects of interventions

Holger J. Schünemann, Peter Tugwell, Barnaby C. Reeves, Elie A. Akl, Nancy Santesso, Frederick A. Spencer, Beverley Shea, George Wells and Mark Helfand
Table 5a. Example of presentation for judgements about indirectness (not to be used for decision making). In people with cancer, does treatment with heparins compared to no treatment reduce mortality (other outcomes would be increased risk of bleeding, venous thromboembolism and others (Akl, Gunukula et al., 2011)).

<table>
<thead>
<tr>
<th>Domain (original question asked)</th>
<th>Description (evidence found and included, including evidence from other studies) - consider the domains of study design and study execution, inconsistency, imprecision and publication bias</th>
<th>Judgment - Is the evidence is sufficiently direct?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population: All patients with advanced cancer</td>
<td>A total of 8 randomized trials included patients with various types of cancer, 2 trials included only patients with small cell lung cancer, others included predominantly breast cancer. The studies were well executed and enrolled patients that were similar to those seen in practice. There was some degree of inconsistency in the baseline risk and related imprecision. Publication bias was not of concern.</td>
<td>Yes ☑️</td>
</tr>
<tr>
<td>Intervention: Heparins</td>
<td>Trials included both low molecular heparin and unfractionated heparin. The observational studies do not suggest differential effects for the heparins.</td>
<td>Yes ☑️</td>
</tr>
<tr>
<td>Comparator: No anticoagulation</td>
<td>Trials used placebo injections</td>
<td>Yes ☑️</td>
</tr>
<tr>
<td>Direct comparison</td>
<td>Studies directly compared the intervention against the comparator of interest (default)</td>
<td>Yes ☑️</td>
</tr>
<tr>
<td>Outcome: Mortality</td>
<td>Mortality was determined through follow-up of patients in the trial (e.g. telephone)</td>
<td>Yes ☑️</td>
</tr>
<tr>
<td>Final judgment about indirectness across domains for the outcome mortality:</td>
<td>The identified evidence is directly relevant to the question. NRS will not provide strong complimentary data for the effects of the intervention. NRS suggest that the baseline risk for the population is similar in the trials compared to the population not included in trials.</td>
<td>☑️ No indirectness ☐ Serious indirectness ☐ Very serious indirectness</td>
</tr>
</tbody>
</table>

Footnote: The degree of indirectness does not lower our confidence that the estimates of effect would be similar for healthcare decision making. It is not useful to look for NRS evidence.
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.†

<table>
<thead>
<tr>
<th>Outcome after 12 Months</th>
<th>Participants</th>
<th>Relative Risk (95% CI)</th>
<th>Anticipated Absolute Effect</th>
<th>Quality of Evidence (GRADE) and Comments</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>
<td>no. of events per 1000 patients</td>
</tr>
<tr>
<td>Death</td>
<td>6245 (10)</td>
<td>0.94 (0.88–1.00)</td>
<td>501</td>
<td>30 fewer (60 fewer to 0 more)</td>
</tr>
<tr>
<td>Symptomatic VTE</td>
<td>5979 (9)</td>
<td>0.58 (0.50–0.67)</td>
<td>105</td>
<td>5 fewer to 0 more</td>
</tr>
<tr>
<td>Major bleeding</td>
<td>6518 (17)</td>
<td>0.59 (0.54–0.64)</td>
<td>101</td>
<td>5 fewer to 0 more</td>
</tr>
<tr>
<td>Minor bleeding</td>
<td>6020 (9)</td>
<td>0.88 (0.74–1.04)</td>
<td>105</td>
<td>3 fewer to 15 more</td>
</tr>
</tbody>
</table>

Combination of judgments: Publication bias and imprecision

## Assessing Quality of Evidence by Outcome

### 1. Establish initial level of confidence

<table>
<thead>
<tr>
<th>Study design</th>
<th>Initial confidence in an estimate of effect</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 confidence

**Reasons for considering lowering or raising confidence**

- **↓ Lower if**
  - Risk of Bias
  - Inconsistency
  - Indirectness
  - Imprecision
  - Publication bias

- **↑ Higher if**
  - Large effect
  - Dose response
  - All plausible confounding & bias
    - would reduce a demonstrated effect or
    - would suggest a spurious effect if no effect was observed

### 3. Final level of confidence rating

- **High**
  - (5 of 5)
- **Moderate**
  - (4 of 5)
- **Low**
  - (3 of 5)
- **Very low**
  - (2 of 5)

*upgrading criteria are usually applicable to observational studies only.*

---

*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)*
# 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

<table>
<thead>
<tr>
<th>Study Design</th>
<th>Initial confidence in estimate of effect</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 confidence

<table>
<thead>
<tr>
<th>Reason for considering lowering or raising confidence</th>
<th>Lowered confidence</th>
<th>Raised confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk of bias</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inconsistency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indirectness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imprecision</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Publication bias</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 3. Final level of confidence rating

<table>
<thead>
<tr>
<th>Confidence in estimate of effect across those considerations</th>
<th>High</th>
<th>Moderate</th>
<th>Low</th>
<th>Very Low</th>
</tr>
</thead>
</table>

*upgrading criteria are usually applicable to observational studies only.*
Altering confidence in observational studies

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

For each outcome based on 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

<table>
<thead>
<tr>
<th>Study Design</th>
<th>Initial confidence in estimate of effect</th>
<th>Reasons for considering lowering or raising confidence</th>
<th>Confidence in estimate of effect across those considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized trials</td>
<td>High confidence</td>
<td>Risk of Bias, Inconsistency, Indirectness, Imprecision</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>Low confidence</td>
<td>Publication bias</td>
<td>Moderate</td>
</tr>
<tr>
<td>Observational studies</td>
<td>Low confidence</td>
<td></td>
<td>Low</td>
</tr>
</tbody>
</table>

*Upgrading criteria are usually applicable to observational studies only.*
Determinants of confidence 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
What can raise quality?

1. large magnitude can upgrade (RRR 50%/RR 2)
   – very large two levels (RRR 80%/RR 5)
   – criteria
     • everyone used to do badly
     • almost everyone does well
   – parachutes to prevent death when jumping from airplanes
Reminders for immunization uptake

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Patient Reminder Sum</th>
<th>Control</th>
<th>Odds Ratio M-H,Random,95% CI</th>
<th>Weight</th>
<th>Odds Ratio H,Random,95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 Other-adult</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hogg1998T101</td>
<td>21/866</td>
<td>4/458</td>
<td></td>
<td>0.9 %</td>
<td>2.82 [ 0.96, 8.27 ]</td>
</tr>
<tr>
<td>Sansom2003T514</td>
<td>242/279</td>
<td>197/245</td>
<td></td>
<td>2.7 %</td>
<td>1.59 [ 1.00, 2.55 ]</td>
</tr>
<tr>
<td>Siebers1985T36</td>
<td>20/72</td>
<td>3/39</td>
<td></td>
<td>0.6 %</td>
<td>4.62 [ 1.28, 16.70 ]</td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td><strong>1217</strong></td>
<td><strong>742</strong></td>
<td></td>
<td>4.2 %</td>
<td>2.19 [ 1.21, 3.99 ]</td>
</tr>
</tbody>
</table>

Total events: 283 (Patient Reminder Sum), 204 (Control)
Heterogeneity: Tau² = 0.10; Chi² = 2.93, df = 2 (P = 0.23); I² = 32%
Test for overall effect: Z = 2.57 (P = 0.010)

Citation: Jacobson Vann JC, Szilagyi P. Patient reminder and recall systems to improve immunization rates. Cochrane Database of Systematic Reviews 2005, Issue 3. Art. No.: CD003941. DOI: 10.1002/14651858.CD003941.pub2.
What can raise quality?

2. dose response relation
   - (higher INR – increased bleeding)
   - childhood lymphoblastic leukemia
     • risk for CNS malignancies 15 years after cranial irradiation
       • no radiation: 1% (95% CI 0% to 2.1%)
       • 12 Gy: 1.6% (95% CI 0% to 3.4%)
       • 18 Gy: 3.3% (95% CI 0.9% to 5.6%)

3. all plausible residual confounding or bias may be working to reduce the demonstrated effect or increase the effect if no effect was observed
All plausible residual confounding and bias would result in an overestimate of effect

- Hypoglycaemic drug phenformin causes lactic acidosis
- The related agent metformin is under suspicion for the same toxicity.
- Large observational studies have failed to demonstrate an association
  - Clinicians would be more alert to lactic acidosis in the presence of the agent
- Vaccine – adverse effects
Evidence profiles

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

### Quality assessment

<table>
<thead>
<tr>
<th>Participants (studies)</th>
<th>Risk of bias</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Publication bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
<td>no serious</td>
<td>no serious</td>
<td>no serious</td>
<td>no serious</td>
<td>undetected</td>
</tr>
<tr>
<td>(3 studies)</td>
<td>risk of bias</td>
<td>inconsistency</td>
<td>indirectness</td>
<td>imprecision</td>
<td></td>
</tr>
<tr>
<td>Hospitalisation</td>
<td>no serious</td>
<td>no serious</td>
<td>no serious</td>
<td>no serious</td>
<td>undetected</td>
</tr>
<tr>
<td>(5 studies)</td>
<td>risk of bias</td>
<td>inconsistency</td>
<td>indirectness</td>
<td>imprecision</td>
<td></td>
</tr>
</tbody>
</table>

### Summary of Findings

#### Study event rates (%)

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>log[Odds Ratio]</th>
<th>SE</th>
<th>Total</th>
<th>Total</th>
<th>Weight</th>
<th>Odds Ratio 95% CI</th>
<th>Odds Ratio 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hanshaoworakul 2009</td>
<td>-2.040221</td>
<td>0.58739416</td>
<td>315</td>
<td>310</td>
<td>28.5%</td>
<td>0.13 [0.04, 0.41]</td>
<td></td>
</tr>
<tr>
<td>Liem 2009 (1)</td>
<td>-0.941609</td>
<td>0.75113239</td>
<td>55</td>
<td>12</td>
<td>17.5%</td>
<td>0.39 [0.09, 1.70]</td>
<td></td>
</tr>
<tr>
<td>McGeer 2009 (2)</td>
<td>-1.309333</td>
<td>0.4270348</td>
<td>69</td>
<td>100</td>
<td>54.0%</td>
<td>0.27 [0.12, 0.62]</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>-2.040221</td>
<td>0.58739416</td>
<td>439</td>
<td>242</td>
<td>100.0%</td>
<td>0.23 [0.13, 0.43]</td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: Tau^2 = 0.00; Chi^2 = 1.58, df = 2 (P = 0.45); I^2 = 0%

Test for overall effect: Z = 4.63 (P < 0.00001)

(1) Adjusted for neutropenia and hospital admission
(2) Does not specify what was adjusted for

#### Complications - Pneumonia

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>log[Odds Ratio]</th>
<th>SE</th>
<th>Total</th>
<th>Total</th>
<th>Weight</th>
<th>Odds Ratio 95% CI</th>
<th>Odds Ratio 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hanshaoworakul 2009</td>
<td>2.040221</td>
<td>0.58739416</td>
<td>2111</td>
<td>100449</td>
<td>21.1%</td>
<td>0.83 [0.59 to 1.16]</td>
<td></td>
</tr>
<tr>
<td>Liem 2009 (1)</td>
<td>0.941609</td>
<td>0.75113239</td>
<td>647</td>
<td>50017</td>
<td>1.3%</td>
<td>0.83 [0.59 to 1.16]</td>
<td></td>
</tr>
<tr>
<td>McGeer 2009 (2)</td>
<td>1.309333</td>
<td>0.4270348</td>
<td>21</td>
<td>21</td>
<td>100.0%</td>
<td>0.83 [0.59 to 1.16]</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2.040221</td>
<td>0.58739416</td>
<td>2111</td>
<td>100449</td>
<td>21.1%</td>
<td>0.83 [0.59 to 1.16]</td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: Tau^2 = 0.00; Chi^2 = 1.58, df = 2 (P = 0.45); I^2 = 0%

Test for overall effect: Z = 4.63 (P < 0.00001)

(1) Adjusted for neutropenia and hospital admission
(2) Does not specify what was adjusted for

(3) Anticipated absolute effects

<table>
<thead>
<tr>
<th>Event</th>
<th>Study event rates (% of total)</th>
<th>Relative effect (adj OR) (95% CI)</th>
<th>Anticipated absolute effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
<td>681 (3 studies)</td>
<td>no serious risk of bias</td>
<td>-200/1032 (19.4%)</td>
</tr>
<tr>
<td>Hospitalisation</td>
<td>150710 (5 studies)</td>
<td>no serious risk of bias</td>
<td>-647/50017 (1.3%)</td>
</tr>
<tr>
<td>Complications - Pneumonia</td>
<td>150466 (3 studies)</td>
<td>no serious risk of bias</td>
<td>-2111/100449 (2.1%)</td>
</tr>
</tbody>
</table>

Hsu J, Santesso N, Mustafa R, et al.

*Antivirals for treatment of influenza: a systematic review and meta-analysis of observational studies.*

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

<table>
<thead>
<tr>
<th>Quality assessment</th>
<th>Summary of Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants (studies)</td>
<td>Risk of bias</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Mortality</td>
<td>no serious risk of bias</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospitalisation</td>
<td>no serious risk of bias</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>ICU admissions/mechanical ventilation/respiratory failure</td>
<td>Serious⁵</td>
</tr>
<tr>
<td>Complications - Pneumonia</td>
<td>no serious risk of bias</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reliable?

<table>
<thead>
<tr>
<th></th>
<th>GRADE working group members (N=15) Reliability coefficient (95% CI)</th>
<th>HRM students (N=10) Reliability coefficient (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GRADE</strong></td>
<td>0.72 (0.61-0.79)</td>
<td>0.66 (0.56-0.75)</td>
</tr>
<tr>
<td><strong>Visual analogue scale without utilizing GRADE</strong></td>
<td>0.27 (0.18-0.37)</td>
<td>0.31 (0.21-0.42)</td>
</tr>
</tbody>
</table>
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.**

<table>
<thead>
<tr>
<th>Outcome after 12 Months</th>
<th>Participants</th>
<th>Relative Risk (95% CI)</th>
<th>Anticipated Absolute Effect</th>
<th>Quality of Evidence (GRADE) and Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Risk without LMWH (95% CI)</td>
<td>Risk Difference with LMWH (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>no. of events per 1000 patients</td>
<td></td>
</tr>
<tr>
<td>Major bleeding</td>
<td>6518 (11)</td>
<td>1.06 (0.71–1.57)</td>
<td>16</td>
<td>Moderate-quality evidence owing to imprecision; the increase may be acceptable to patients, given that VTE, which occurs more frequently, may be equally unpleasant</td>
</tr>
<tr>
<td>Minor bleeding</td>
<td>6020 (9)</td>
<td>1.18 (0.89–1.55)</td>
<td>27</td>
<td>Moderate-quality evidence owing to imprecision; however, this outcome is unlikely to be critical for decision making</td>
</tr>
</tbody>
</table>

Should every cancer patient receive heparin?
Balancing desirable and undesirable consequences

For

Against

Effects & $x$
values

Effects & $x$
values

Conditional

Strong
From evidence to recommendations

RCTs

Observational studies

High level recommendation

Lower level recommendation

Balance between benefits, harms & burdens

Quality of evidence

Patients’ values & preferences

Resource use

GRADE

Old systems

GRADE
Should heparin be used in patients with cancer who have no other therapeutic or prophylactic indication for anticoagulation?

<table>
<thead>
<tr>
<th>DOMAIN</th>
<th>JUDGEMENT</th>
<th>DETAILS OF JUDGEMENT</th>
<th>EVIDENCE/EXPLANATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the overall quality of evidence?</td>
<td>High, Moderate, Low, Very low</td>
<td>Consider: QoE for benefits: harms.</td>
<td>This will be critical and important outcomes: mortality, etc.</td>
</tr>
<tr>
<td>Is there high or moderate quality evidence?</td>
<td>High, Moderate, Low, Very low</td>
<td>QoE for benefits: harms.</td>
<td>This is the QoE for values and prefs – see values and prefs evidence tables</td>
</tr>
<tr>
<td>The higher the quality of evidence, the more likely is a strong recommendation</td>
<td>High, Moderate, Low, Very low</td>
<td>QoE for benefits: harms.</td>
<td>Evidence to recommendation/decision framework</td>
</tr>
<tr>
<td>What is the balance between benefits and risks/burden?</td>
<td>High, Moderate, Low, Very low</td>
<td>Baseline risk for benefits, harms and burden:</td>
<td>Of interest is balancing the magnitude of benefits and harms</td>
</tr>
<tr>
<td>Are you confident that the benefits outweigh the harms and burden or vice versa?</td>
<td>High, Moderate, Low, Very low</td>
<td>Baseline risk for benefits, harms and burden:</td>
<td>Of interest is determining values satisfactory for amendment:</td>
</tr>
<tr>
<td>The larger the difference between the benefits and harms, the more likely is a strong recommendation. The smaller the net benefit or harm and the lower the certainty for that net effect, the more likely is a conditional/weak recommendation.</td>
<td>High, Moderate, Low, Very low</td>
<td>Baseline risk for benefits, harms and burden:</td>
<td>Of interest is determining values satisfactory for amendment:</td>
</tr>
<tr>
<td>What are the patient’s values and preferences?</td>
<td>High, Moderate, Low, Very low</td>
<td>Baseline risk for benefits, harms and burden:</td>
<td>Of interest is determining values satisfactory for amendment:</td>
</tr>
<tr>
<td>Are the assumed or identified relative values similar across the largest population?</td>
<td>High, Moderate, Low, Very low</td>
<td>Baseline risk for benefits, harms and burden:</td>
<td>Of interest is determining values satisfactory for amendment:</td>
</tr>
<tr>
<td>The greater the similarity in values and preferences, the more likely is a strong recommendation.</td>
<td>High, Moderate, Low, Very low</td>
<td>Baseline risk for benefits, harms and burden:</td>
<td>Of interest is determining values satisfactory for amendment:</td>
</tr>
</tbody>
</table>

Is there likely large variation seen in values/preferences/utilities
<table>
<thead>
<tr>
<th>Is the incremental cost (or resource use) small relative to the benefits? Are the resources worth the expected net benefit from following the recommendation? The lower the cost of an intervention compared to the alternative, and other costs related to the decision – that is, the fewer resources consumed – the more likely is a strong recommendation in favour of that intervention.</th>
<th>What are the cost per resource unit? Feasibility: Is this intervention generally available? Opportunity cost: Is this intervention and its effects worth withdrawing or not allocating resources from other interventions Differences across settings: Is there lots of variability in resource requirements across settings?</th>
</tr>
</thead>
<tbody>
<tr>
<td>□ Cost is very small relative to the net benefits □ Cost is small relative to the net benefits □ Cost is borderline relative to the net benefits □ Cost is high relative to the net benefits □ Cost is very high relative to the net benefits</td>
<td></td>
</tr>
<tr>
<td>Is there likely large variation seen in values/preferences/utilities</td>
<td>Will the intervention reduce inequities?</td>
</tr>
<tr>
<td>□ High □ Moderate □ Low □ Very low □ Uncertain</td>
<td></td>
</tr>
</tbody>
</table>

If we provide this drug will we not be able to provide drugs for other diseases? Is this drug accessible to all people so would not create inequities?
# Recommendation

<table>
<thead>
<tr>
<th>Recommendation</th>
<th>Should</th>
<th>be used instead of</th>
<th>for</th>
<th>?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall balance of consequences</td>
<td>Undesirable consequences clearly outweigh desirable consequences</td>
<td>Undesirable consequences probably outweigh desirable consequences</td>
<td>The balance between desirable and undesirable consequences is too uncertain*</td>
<td>The balance of desirable and undesirable consequences indicates they are very similar*</td>
</tr>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td></td>
<td>We recommend against the option or for the alternative</td>
<td>We suggest not to use the option or to use the alternative</td>
<td>No recommendation</td>
<td>We suggest using the option</td>
</tr>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

Panel decisions

You can document deliberations of the panel here

Recommendation

Insert the recommendation text here

Remarks and justifications

Explanation

Implementation and feasibility

* In this situation no recommendation could be reasonable
Should heparin be used in patients with cancer who have no other therapeutic or prophylactic indication for

**Population:** Patients with advanced cancer, without other therapeutic or prophylactic indication for anticoagulation

**Intervention:** Heparin

**Comparison:** No anticoagulation

**Setting:** Outpatient

**Perspective:** Health system (*might not be applicable from a individual decision making perspective*)

<table>
<thead>
<tr>
<th>DOMAIN</th>
<th>JUDGEMENT</th>
<th>Research evidence</th>
<th>EXPLANATION/ADDITIONAL INFORMATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the overall quality of evidence?</td>
<td>Is there high or moderate quality evidence?</td>
<td>The higher the quality of evidence, the more likely is a strong recommendation.</td>
<td>There is moderate quality evidence that, at one year, treatment with heparin does not reduce the risk of death and does not increase the risk of major bleeding.</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>High</td>
<td>Somewhat agree</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td></td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Very low</td>
<td></td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td></td>
<td>Disagree</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td></td>
<td>Agree</td>
</tr>
<tr>
<td>What is the balance between benefits and risks/burden?</td>
<td>The larger the difference between the benefits and harms, the more likely is a strong recommendation. The smaller the net benefit or net harm and the lower the certainty for that net effect, the more likely is a conditional/weak recommendation.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Benefits outweigh harms/burden</td>
<td>Benefits slightly outweigh harms/burden</td>
<td>Benefits and harms/burden are balanced</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Participants (studies)</td>
<td>Quality of the evidence</td>
<td>Relative effect (95% CI)</td>
</tr>
<tr>
<td>Death Follow-up 12 months</td>
<td>2,531 (8)</td>
<td>MODERATE</td>
<td>RR 0.93 (0.85 to 1.12)</td>
</tr>
<tr>
<td>Symptomatic VTE Follow-up 12 months</td>
<td>2264 (7)</td>
<td>HIGH</td>
<td>RR 0.55 (0.37 to 0.82)</td>
</tr>
<tr>
<td>Major bleeding Follow-up 12 months</td>
<td>2,843 (9)</td>
<td>MODERATE</td>
<td>RR 1.30 (0.59 to 2.88)</td>
</tr>
</tbody>
</table>

Heparin reduces the risk of VTE, from 5 fewer to 8 fewer per 1000.

Heparin does not reduce the risk of death and does not increase the risk of major bleeding.
**Outcome** | **Quality of life / utility values (range)**
--- | ---
Death | 0
Deep venous thromboembolism | 0.58 (0 to 1.0)
Pulmonary embolism | 0.20 (0 to 1.0)
Nonfatal major extracranial bleeds | 0.8 (0.5 to 0.99)
Minor bleeding | 0.84 (0 to 1.0)
Bleed for low molecular weight heparin | 0.96

*Estimated value of the quality of life with each outcome, where 0.00 represents death and 1.00 represents perfect health.

Quality-adjusted life expectancy are unavailable. Given the likely reduction in hospitalizations related to VTE, and the lack of large effects on adverse effects with limited discomfort from injections, a small net QALY gain can be expected.

No study evaluated values and preferences of the intervention in the target population. Only one small randomized trial (n=138) assessed quality of life as outcome; results were similar among intervention and control group. Source: Sidras K (2)

**Is the incremental cost (or resource use) small relative to the benefits?**
Are the resources worth the expected net benefit from following the recommendation?
The lower the cost of an intervention compared to the alternative, and other costs related to the decision – that is, the fewer resources consumed – the more likely is a strong recommendation in favour of that intervention.

- Cost is very small relative to the net benefits
- Cost is small relative to the net benefits
- Cost is borderline relative to the net benefits
- Cost is high relative to the net benefits
- Cost is very high relative to the net benefits

Formal cost-effectiveness analyses for LMWH for this indication are unavailable. It is assumed that 13 fewer hospitalizations of 2 days on average are avoided by VTE prevention hospitalization in the group without heparin; 29 per 1,000.

<table>
<thead>
<tr>
<th>Total cost for 1,000 patients per year (2008 US $)</th>
<th>No heparin</th>
<th>LMWH</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drug acquisition cost</td>
<td>$0</td>
<td>$9.48</td>
<td>$9.48</td>
</tr>
<tr>
<td>Hospitalizations</td>
<td>$655,078</td>
<td>$35,964</td>
<td>$619,114</td>
</tr>
</tbody>
</table>

The cost to prevent one hospitalization related to VTE is US$ 730,000.

Drug costs are US wholesale acquisition cost data from 2008 for dalteparin. The cost of hospitalizations and investigations for VTE are estimated as $1,122 per day. Source: Candrilli MS (3) The calculation does not include treatment-related serious adverse event costs (e.g., thrombocytopenia) and minor bleeding. The US acquisition cost for dalteparin are for a dose of 40 mg once daily = $25.97, $9,479 per patient per year (2009).

**What would be the impact on health inequities?**

- High
- Moderate
- Low
- Very low
- Uncertain

None

**Is the option feasible to implement?**

- Yes
- Probably yes
- Uncertain
- Probably no
- No

It requires initial investment, and training for patients.
### Recommendation

**Should heparins be used instead of no heparins for patients with advanced cancer?**

<table>
<thead>
<tr>
<th>Overall balance of consequences</th>
<th>Undesirable consequences clearly outweigh desirable consequences</th>
<th>Undesirable consequences probably outweigh desirable consequences</th>
<th>The balance between desirable and undesirable consequences is too uncertain</th>
<th>The balance of desirable and undesirable consequences indicates they are very similar</th>
<th>Desirable consequences probably outweigh undesirable consequences</th>
<th>Desirable consequences clearly outweigh undesirable consequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td>We recommend against the option or for the alternative</td>
<td>We suggest not to use the option or to use the alternative</td>
<td>No recommendation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Panel decisions**

*Despite the cost the benefits outweigh the harms.*

**Recommendation**

*The panel suggests using heparin in patients with advanced solid cancers (weak recommendation, moderate quality evidence)*

**Remarks and justifications**

Values and preferences will differ across patients and shared decision making is most appropriate in this situation.

**Explanation**

The panel assumed that most patients

**Implementation and feasibility**

Resource implications are high in regards to drug acquisition but in view of resource utilization for other cancer therapies and avoidance of hospitalizations possibly justified. Patients or relatives will have to learn injection techniques.

- In this situation no recommendation could be reasonable
Guidelines and decision making

Clinical Practice Guideline

Decision points:

Low uncertainty / Strong recommendation (e.g. aspirin use in myocardial infarction)

Supporting optimal behaviors

INFORMATION COMPONENTS
- Define clear recommendation
- Communicate benefits and risks to explain the rationale

BEHAVIOR CHANGE COMPONENTS
- Implementation strategies
- Performance measures based on professional/patient behavior (prescribing aspirins/taking aspirins)

High uncertainty / Conditional recommendation (e.g. lumpectomy vs mastectomy in breast ca)

Supporting deliberation

INFORMATION COMPONENTS
- Make options explicit
- Communicate benefits and risks of options to explain the dilemma

DECISION MAKING COMPONENTS
- Deliberation methods
- Preference constructing methods
- Performance measures based on quality of decision process (e.g. use of breast cancer decision aid)

v.d. Weijden J Clin Epi 2012
Formulate question
Select outcomes	Rate importance

Outcomes across studies
Create evidence profile with GRADEpro

Rate quality of evidence for each outcome

Randomization increases initial quality

Evidence synthesis

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

Grade down
Grade up

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

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

Evidence synthesis

Outcome Critical
Outcome Critical
Outcome Important
Outcome Not important

Summary of findings & estimate of effect for each outcome

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)

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

Guideline

♦♦♦♦
♦♦♦
♦♦
♦√
√√
√√√
√√√√

♦♦♦♦♦
♦♦♦♦♦
♦♦♦♦♦
♦♦♦♦♦
♦♦♦♦♦
Background & update: GRADE

- Focused on public health, health policy, prognosis and diagnostic questions and how to use evidence to make recommendations (in guidelines or other health care context).

**Recommendation**

2.1.10. For patients with AF, including those with paroxysmal AF, who are at high risk of stroke (e.g., CHADS² score = 2), we recommend oral anticoagulation rather than no therapy (Grade 1A).

2.1.4. In patients with acute DVT of the leg, we suggest early ambulation over initial bed rest (Grade 2C).

- DECIDE research project (www.decide-collaboration.eu)
  - EU FP7

Treweek et al. Impl Science 2013
Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence

Ten partners in seven countries (8 workpackages) EU FP7
University of Dundee, UK
Norwegian Knowledge Centre for the Health Services, Norway
Research Institute of the Hospital of Santa Creu and Sant Pau, Spain
Italian Cochrane Centre, Italy
University of Amsterdam, the Netherlands
World Health Organisation, International
Freiburg University Hospital, Germany
National Institute for Health and Clinical Excellence, UK
Scottish Intercollegiate Guidelines Network, UK
Finnish Medical Society Duodecim, Finland
DECIDE Work Package Table

WP8 Project Management

WP7 Dissemination and Exploitation

- Strategies for targeted dissemination to key stakeholders who determine what happens in clinical practice
  - WP1: Strategies for communicating evidence-based recommendations to clinicians
  - WP2: Strategies for communicating evidence-based recommendations to policymakers and managers
  - WP3: Strategies for communicating evidence-based recommendations to patients and the general public

- Strategies for targeted dissemination of different types of recommendations
  - WP4: Strategies for communicating evidence-based recommendations about diagnostic tests
  - WP5: Strategies for communicating evidence-based recommendations about public health and health systems policies

WP6: Strategies for collaboration among guideline developers and HTA agencies in Europe
WP 6: Strategies for collaboration among European guideline developers and health technology assessment agencies in Europe

The Guideline Development Tool (GDT) is a software bundle supporting the process of development of health care recommendations (e.g., in clinical practice guidelines). Integral to its use will be the completely redesigned GRADEpro – software used by authors of systematic reviews, guideline developers and others involved in health care decision making to create summaries of the evidence supporting health care decisions. The complete bundle will provide additional modules facilitating various stages of summarizing of the evidence and the process of decision making in health care, e.g., clinical and public health guidelines and coverage decisions.

www.guidelinedevelopment.org
The modules will include among others:

- Topic proposal and selection
- Developing the scope
- Generation of structured health care questions (PICO)
- Identifying and rating importance of outcomes
- Evidence retrieval
- Designing search strategies
- Reference management
- Data extraction and management
- GRADEprofiler
- Word processor
- Decision support tool
- Research tools

Copyright © 2012, McMaster University. All rights reserved.
The development of GRADEprofiler (GRADEpro) has been partially supported from the European Union Seventh Framework Programme (FP7 – HEALTH.2010.3.1-1 – two stage) under grant agreement n° 255583.
Web-based tool
Parenteral anticoagulation in patients with cancer who have no therapeutic or prophylactic indication for anticoagulation
Parenteral anticoagulation in patients with cancer who have no therapeutic or prophylactic indication for anticoagulation
Parenteral anticoagulation in patients with cancer who have no therapeutic or prophylactic indication for anticoagulation

| Title: | Provide title of the document/guideline that will reflect its scope. |
| Purpose: | Specify health intents (i.e., prevention, diagnosis, treatment, etc.) and expected benefits or outcomes. E.g., preventing thromboembolic complications of patients undergoing elective orthopedic surgery. |
| Target population: | Specify subjects to whom those recommendations apply (i.e., patients, society, etc.) E.g., adults undergoing elective orthopedic surgery, all women 40 years of age or older, etc. |
| Healthcare setting: | Specify level of health care (i.e., primary, secondary, etc.) where these recommendations are supposed to be implemented. |
| Types of interventions: | Specify which preventive, therapeutic and diagnostic interventions will be covered and which will not. |
| Key stakeholders and users: | Specify all relevant professional groups, institutions, patients, public, etc. who are target users or beneficiaries of these guidelines and/or whose views should be sought. |
| Key resources to consider: | Specify resources needed for the implementation of guidelines (i.e., need for additional human resources, equipment, infrastructure, system changes, etc.) and potential barriers to implementation. |
| Existing documents: | List all existing documents/guidelines on the same or similar topic that are likely to be currently used in practice (e.g., guidelines developed by other organizations). |
Parenteral anticoagulation in patients with cancer who have no therapeutic or prophylactic indication for anticoagulation

[DOCUMENT TITLE]

[DOCUMENT SUBTITLE]

Authors

Disclosure of potential conflicts of interest:

Review group

Table of contents
Parenteral anticoagulation in patients with cancer who have no therapeutic or prophylactic indication for anticoagulation

Should Heparin vs placebo be used for patients with cancer who have no therapeutic or prophylactic indication for anticoagulation?

- Add management question
- Add diagnostic question
Parenteral anticoagulation in patients with cancer who have no therapeutic or prophylactic indication for anticoagulation

### Should Heparin vs placebo be used for patients with cancer who have no therapeutic or prophylactic indication for anticoagulation?

<table>
<thead>
<tr>
<th>Quality assessment</th>
<th>No of patients</th>
<th>Summary of findings</th>
<th>Effect</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No of studies</td>
<td>Design studies</td>
<td>Risk of bias</td>
<td>Inconsistency</td>
<td>Indirectness</td>
</tr>
<tr>
<td>Mortality at 12 months</td>
<td>9</td>
<td>1433/3072 (46.6)%</td>
<td>1308/2670 (49.0)%</td>
<td>RR 0.94 (0.88 to 1.01)</td>
<td>29 fewer per 1000 (from 5 more to 59 fewer)</td>
</tr>
<tr>
<td>Mortality at 24 months</td>
<td>5</td>
<td>461/586 (78.7)%</td>
<td>505/588 (85.9)%</td>
<td>RR 0.92 (0.88 to 0.97)</td>
<td>69 fewer per 1000 (from 26 fewer to 103 fewer)</td>
</tr>
<tr>
<td>Mortality over duration of study</td>
<td>8</td>
<td>0/2305 (0.0)%</td>
<td>0/2290 (0.0)%</td>
<td>HR 0.82 (0.70 to 0.95)</td>
<td>0 fewer per 1000 (from 0 fewer to 0 fewer)</td>
</tr>
<tr>
<td>Symptomatic VTE</td>
<td>8</td>
<td>58/2946 (2.0)%</td>
<td>112/2530 (4.4)%</td>
<td>RR 0.47 (0.34 to 0.64)</td>
<td>23 fewer per 1000 (from 16 fewer to 29 fewer)</td>
</tr>
<tr>
<td>Major bleeding</td>
<td>10</td>
<td>49/3213 (1.5)%</td>
<td>41/2802 (1.5)%</td>
<td>RR 1.08 (0.65 to 1.82)</td>
<td>1 more per 1000 (from 5 fewer to 12 more)</td>
</tr>
<tr>
<td>Minor bleeding</td>
<td>8</td>
<td>111/2954 (3.8)%</td>
<td>64/2363 (2.5)%</td>
<td>RR 1.23 (0.86 to 1.74)</td>
<td>6 more per 1000 (from 3 fewer to 18 more)</td>
</tr>
</tbody>
</table>

Add outcome

Footnotes
<table>
<thead>
<tr>
<th>Outcome</th>
<th>No of participants (Studies)</th>
<th>Evidence tables</th>
<th>Relative effect (95% CI)</th>
<th>Anticipated absolute effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality at 12 months</td>
<td>5742 (9 Studies)</td>
<td></td>
<td>RR 0.94 (0.88 to 1.01)</td>
<td>490 per 1000</td>
</tr>
<tr>
<td>Mortality at 24 months</td>
<td>1174 (6 Studies)</td>
<td></td>
<td>RR 0.92 (0.88 to 0.97)</td>
<td>859 per 1000</td>
</tr>
<tr>
<td>Mortality over duration of study</td>
<td>4593 (8 Studies)</td>
<td></td>
<td>HR 0.82 (0.70 to 0.95)</td>
<td>0 per 1000</td>
</tr>
<tr>
<td>Symptomatic VTE</td>
<td>5476 (8 Studies)</td>
<td></td>
<td>RR 0.47 (0.34 to 0.64)</td>
<td>44 per 1000</td>
</tr>
<tr>
<td>Major bleeding</td>
<td>6015 (10 Studies)</td>
<td></td>
<td>RR 1.08 (0.65 to 1.82)</td>
<td>15 per 1000</td>
</tr>
<tr>
<td>Minor bleeding</td>
<td>5517 (8 Studies)</td>
<td></td>
<td>RR 1.23 (0.86 to 1.74)</td>
<td>25 per 1000</td>
</tr>
</tbody>
</table>
**Parenteral anticoagulation in patients with cancer who have no therapeutic or prophylactic indication for anticoagulation**

**Should Heparin vs placebo be used for patients with cancer who have no therapeutic or prophylactic indication for anticoagulation?**

<table>
<thead>
<tr>
<th>No of studies</th>
<th>Design studies</th>
<th>Risk of bias</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
<th>Quality assessment</th>
<th>No of patients</th>
<th>Summary of findings</th>
<th>Effect</th>
<th>Absolute</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1433/5072 (44.6%)</td>
<td>1309/2670 (49.0%)</td>
<td>RR 0.94 (0.88 to 1.01)</td>
<td>29 fewer per 1000 (from 5 more to 59 fewer)</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>461/586 (78.7%)</td>
<td>501/588 (85.9%)</td>
<td>RR 0.92 (0.88 to 0.97)</td>
<td>69 fewer per 1000 (from 26 fewer to 103 fewer)</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0/2303 (0.0%)</td>
<td>0/2290 (0.0%)</td>
<td>HR 0.82 (0.70 to 0.95)</td>
<td>0 fewer per 1000 (from 0 fewer to 0 fewer)</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>58/2946 (2.0%)</td>
<td>112/2330 (4.4%)</td>
<td>RR 0.47 (0.34 to 0.64)</td>
<td>23 fewer per 1000 (from 16 fewer to 29 fewer)</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>49/3213 (1.5%)</td>
<td>41/2802 (1.5%)</td>
<td>RR 1.08 (0.65 to 1.82)</td>
<td>1 more per 1000 (from 5 fewer to 12 more)</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>111/2954 (3.8%)</td>
<td>64/2563 (2.5%)</td>
<td>RR 1.23 (0.86 to 1.74)</td>
<td>6 more per 1000 (from 3 fewer to 18 more)</td>
</tr>
</tbody>
</table>

**Footnotes**
Should Heparin vs placebo be used for patients with cancer who have no therapeutic or prophylactic indication for anticoagulation?

<table>
<thead>
<tr>
<th>Quality assessment</th>
<th>No of patients</th>
<th>Summary of findings</th>
<th>Effect</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>Relative (95% CI)</td>
<td>Absolute</td>
<td></td>
</tr>
<tr>
<td>Mortality at 12 months</td>
<td></td>
<td>1433/1072 (46.6%)</td>
<td>RR 0.94 (0.88 to 1.01)</td>
<td>29 fewer per 1000 (from 5 more to 59 fewer)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1308/2670 (49.0%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality at 24 months</td>
<td></td>
<td>461/586 (78.7%)</td>
<td>RR 0.92 (0.88 to 0.97)</td>
<td>69 fewer per 1000 (from 26 fewer to 103 fewer)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>503/588 (85.9%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality over duration of study</td>
<td></td>
<td>0/2501 (0.0%)</td>
<td>HR 0.82 (0.70 to 0.95)</td>
<td>0 fewer per 1000 (from 0 fewer to 0 fewer)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0/2290 (0.0%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symptomatic VTE</td>
<td></td>
<td>58/2946 (2.0%)</td>
<td>RR 0.47 (0.34 to 0.64)</td>
<td>23 fewer per 1000 (from 16 fewer to 29 fewer)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>112/2530 (4.4%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major bleeding</td>
<td></td>
<td>49/5213 (1.5%)</td>
<td>RR 1.08 (0.65 to 1.82)</td>
<td>1 more per 1000 (from 5 fewer to 12 more)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>41/2802 (1.5%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minor bleeding</td>
<td></td>
<td>111/2954 (3.8%)</td>
<td>RR 1.23 (0.86 to 1.74)</td>
<td>6 more per 1000 (from 3 fewer to 18 more)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>64/2563 (2.5%)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Footnotes
### Parenteral anticoagulation in patients with cancer who have no therapeutic or prophylactic indication for anticoagulation

#### Outcome: Mortality at 12 months

<table>
<thead>
<tr>
<th>Domain (original question asked)</th>
<th>Description (evidence found and included, including evidence from other studies) - consider the domains of study design and study execution, inconsistency, imprecision and publication bias</th>
<th>Judgment - Is the evidence sufficiently direct?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population:</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Intervention: Heparin</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Comparator: placebo</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Direct comparison</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Outcome: Mortality at 12 months</td>
<td></td>
<td>Yes</td>
</tr>
</tbody>
</table>

Final judgment about Indirectness across domains:

- No Indirectness
- Serious Indirectness
- Very serious Indirectness

[Cancel] [Apply]
Table 5a. Example of presentation for judgements about indirectness (not to be used for decision making). In people with cancer, does treatment with heparins compared to no treatment reduce mortality (other outcomes would be increased risk of bleeding, venous thromboembolism and others (Akl, Gunukula et al., 2011)).

<table>
<thead>
<tr>
<th>Domain (original question asked)</th>
<th>Description (evidence found and included, including evidence from other studies) - consider the domains of study design and study execution, inconsistency, imprecision and publication bias</th>
<th>Judgment - Is the evidence is sufficiently direct?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population: All patients with advanced cancer</td>
<td>A total of 8 randomized trials included patients with various types of cancer, 2 trials included only patients with small cell lung cancer, others included predominantly breast cancer. The studies were well executed and enrolled patients that were similar to those seen in practice. There was some degree of inconsistency in the baseline risk and related imprecision. Publication bias was not of concern.</td>
<td><img src="%E2%98%91" alt="Yes" /> <img src="%E2%98%92" alt="Probably yes" /> <img src="%E2%98%90" alt="Probably no" /> <img src="%E2%98%90" alt="No" /></td>
</tr>
<tr>
<td>Intervention: Heparins</td>
<td>Trials included both low molecular heparin and unfractionated heparin. The observational studies do not suggest differential effects for the heparins.</td>
<td><img src="%E2%98%91" alt="Yes" /> <img src="%E2%98%92" alt="Probably yes" /> <img src="%E2%98%90" alt="Probably no" /> <img src="%E2%98%90" alt="No" /></td>
</tr>
<tr>
<td>Comparator: No anticoagulation</td>
<td>Trials used placebo injections</td>
<td><img src="%E2%98%91" alt="Yes" /> <img src="%E2%98%92" alt="Probably yes" /> <img src="%E2%98%90" alt="Probably no" /> <img src="%E2%98%90" alt="No" /></td>
</tr>
<tr>
<td>Direct comparison</td>
<td>Studies directly compared the intervention against the comparator of interest (default)</td>
<td><img src="%E2%98%91" alt="Yes" /> <img src="%E2%98%92" alt="Probably yes" /> <img src="%E2%98%90" alt="Probably no" /> <img src="%E2%98%90" alt="No" /></td>
</tr>
<tr>
<td>Outcome: Mortality</td>
<td>Mortality was determined through follow-up of patients in the trial (e.g. telephone)</td>
<td><img src="%E2%98%91" alt="Yes" /> <img src="%E2%98%92" alt="Probably yes" /> <img src="%E2%98%90" alt="Probably no" /> <img src="%E2%98%90" alt="No" /></td>
</tr>
<tr>
<td>Final judgment about indirectness across domains for the outcome mortality:</td>
<td>The identified evidence is directly relevant to the question. NRS will not provide strong complimentary data for the effects of the intervention. NRS suggest that the baseline risk for the population is similar in the trials compared to the population not included in trials.</td>
<td><img src="%E2%98%92" alt="No indirectness" /> <img src="%E2%98%90" alt="Serious indirectness" /> <img src="%E2%98%90" alt="Very serious indirectness" /></td>
</tr>
</tbody>
</table>

Footnote: The degree of indirectness does not lower our confidence that the estimates of effect would be similar for healthcare decision making. It is not useful to look for NRS evidence.
Parenteral anticoagulation in patients with cancer who have no therapeutic or prophylactic indication for anticoagulation

**Should Heparin vs placebo be used for patients with cancer who have no therapeutic or prophylactic indication for anticoagulation?**

<table>
<thead>
<tr>
<th>DOMAIN</th>
<th>JUDGEMENT</th>
<th>DETAILS OF JUDGEMENT</th>
<th>EVIDENCE/EXPLANATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>QUALITY</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is the overall quality of evidence?</td>
<td>High, Moderate, Low, Very low</td>
<td>Critical Outcomes: Mortality at 12 months, Mortality at 24 months, Mortality over duration of study, Symptomatic VTE, Major bleeding</td>
<td>Agree, Somewhat agree, Uncertain, Somewhat disagree, Disagree</td>
</tr>
<tr>
<td>What is the balance between benefits and harms/burden?</td>
<td>Benefits outweigh harms/burden, Slightly outweigh harms/burden, Harms/burden are balanced, Slightly outweigh benefits, Harms/burden outweigh benefits</td>
<td>Critical Outcomes: Mortality at 12 months, Mortality at 24 months, Mortality over duration of study, Symptomatic VTE, Major bleeding</td>
<td>Agree, Somewhat agree, Somewhat disagree, Disagree</td>
</tr>
<tr>
<td>What are the patient's values and preferences?</td>
<td>Similar values</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**OUTCOMES**

**SEARCHING**

**SCREENING**

**DATA EXTRACTION**

**RISK OF BIAS**

**ANALYSES**

**EVIDENCE TABLE**

**RECOMMENDATIONS**
GRADE for diagnosis

• Shares fundamental logic of other interventions

• If a test fails to improve important outcomes (incl. reducing resources, complications): no reason to use it, whatever its accuracy

• Recognized for over 20 years

• Best way to assess diagnostic strategy: randomized controlled trial or well-done observational studies in which investigators focus on patient-important outcomes

• However, poor guidance for what to do if no such evidence available
Are there studies that directly focus on: mortality, morbidity, symptoms, and/or quality of life?

Apply GRADE approach as for treatment or other intervention

Schunemann et al. BMJ, 2008
Study designs II

Accuracy Study

Target population

New test(s) + Reference test

TP  FP  FN  TN

Assumptions or indirect evidence about management of patients correctly or incorrectly classified as positive or negative with the new or old test(s)

Judgements about patient-important outcomes with a new test and a reference test

Look for diagnostic test accuracy studies

And then draw inferences from other evidence

Schunemann et al. BMJ, 2008
Diagnostic Test Accuracy Synthesis

Diagnostic Test Accuracy Quality of evidence

5 factors to downgrade (factors to upgrade?)

- High
- Moderate
- Low
- Very low

TP
FP
TN
FN

Apply GRADE for Diagnostic Test Accuracy studies

Create DTA evidence synthesis (pooled estimates of DTA)

Summary of findings & estimate of effect for each outcome

Step 1
Determinants of quality for diagnostic test accuracy

• Observational studies: ⊕⊕⊕⊕⊕

• 5 factors that can lower quality
  1. limitations in detailed design and execution (risk of bias criteria)
  2. Inconsistency (or heterogeneity)
  3. Indirectness (PICO and applicability)
  4. Imprecision (number of events and confidence intervals)
  5. Publication bias

• Theoretically factors can increase quality
  1. large magnitude of effect
  2. dose-response gradient
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.”
**Diagnostic Test Accuracy Synthesis**

**Diagnostic Test Accuracy**

Quality of evidence

- Very low
- Low
- Moderate
- High

⊕⊕⊕⊕
⊕
⊕⊕
OO

**Assess linked evidence**

Directness of the outcome

- Very uncertain
- Uncertain
- Moderately certain
- Certain
- Very certain

Lower our confidence

**Final Quality of evidence for each outcome based on DTA and linked evidence and development of recommendations**

**Mortality**

High ⊕⊕⊕⊕
Moderate ⊕⊕⊕O
Low ⊕⊕OO
Very low ⊕OOO

**Morbidity**

**QoL**

**Harms**

**Resources**

**Other**

Natural History
Patients will suffer from disease without being detected or suffer from symptoms and undergo repeat testing or testing for other disease that will happen at certain rate.

SR required, full framework always needs to be developed.

**Step 2**

**Apply GRADE for Diagnostic Test Accuracy studies**

TP
FP
TN
FN

5 factors to downgrade
(factors to upgrade?)

- High ⊕⊕⊕⊕
- Moderate ⊕⊕⊕O
- Low ⊕⊕OO
- Very low ⊕OOO

**STUDIES**

Sensitivity
Specificity

**Epidemiology**

**Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)**

**Quality of study**

- Poor
- Fair
- Good
- Very good

⊕⊕⊕⊕
⊕
⊕⊕
OO

Apply GRADE for Diagnostic Test Accuracy studies

e.g. 3%

Enter the additional text for the image description here.
Summary

• Complex judgments about evidence
• GRADE has become the international standard
  • Benefits/efficacy
  • Harms/safety
  • Diagnosis
• Transparency and frameworks required to assess quality of a body of evidence
  – 8 criteria in GRADE
• Recommendation depends on transparent judgments about the balance between desirable and undesirable consequences and our confidence in them (in terms of patient/population-important outcomes)
### Telephone counselling to improve adherence to diet

**Narrative synthesis**

Total number of studies: 4  
Total number of participants: 255

<table>
<thead>
<tr>
<th>Study</th>
<th>Statistic/Measure</th>
<th>Results</th>
<th>P value/RR</th>
</tr>
</thead>
</table>
| Chui 2005        | Median (IQR) 0 (none) – 3 (complete adherence) | Telephone (n=31): 1 (0-1)  
Control (n=32): 0 (0-0)  
- “Slight improvement with telephone intervention” | P value (0.32) “calculate effect” |
| Stewart 2005     | Number adhering to diet            | Telephone (n=40): 26/40  
Control (n=38): 15/38  
- “Slight improvement with telephone intervention” | RR 1.65 (1.05, 2.59) |
| Racelis 1998     | Diet score                         | Telephone (n=11): improved  
Control (n=10): improved  
- “no effect of telephone intervention vs control” | “No significant difference” |
| Cummings 1981    | Number compliant                   | All groups (n=93) compliance pre = 86%, post = 90%  
- “study cannot be used” | Not reported |
Risk of bias

<table>
<thead>
<tr>
<th>Study</th>
<th>Random sequence generation (selection bias)</th>
<th>Allocation concealment (selection bias)</th>
<th>Blinding (performance bias and detection bias): Participants</th>
<th>Blinding (performance bias and detection bias): Providers</th>
<th>Blinding (performance bias and detection bias): Outcome assessors</th>
<th>Incomplete outcome data (attrition bias)</th>
<th>Selective reporting (reporting bias)</th>
<th>Other bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stewart 2005</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
• Risk of bias
  – Medium
• Imprecision
  – All together 162 participants with small effect
    • OIS not met
• Inconsistency
  – Yes, some inconsistency
• Publication bias
  – No small negative study?
• Indirectness
  – No concern

Strength of evidence?
Downgrade for risk of bias (-1) & imprecision (-1) & the remaining criteria in context (-1)

very low/insufficient
Telephone counselling to improve adherence to diet

Narrative synthesis
Total number of studies: 4
Total number of participants: 255

<table>
<thead>
<tr>
<th>Study</th>
<th>Statistic/Measure</th>
<th>Results</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chui 2005</td>
<td>Median (IQR) 0 (none) – 3 (complete adherence)</td>
<td>Telephone (n=31): 1 (0-1) Control (n=32): 0 (0-0) - “Slight improvement with telephone intervention”</td>
<td>P value (0.32) “calculate effect”</td>
</tr>
<tr>
<td>Stewart 2005</td>
<td>Number adhering to diet</td>
<td>Telephone (n=40): 26/40 Control (n=38): 15/38 - “Slight improvement with telephone intervention”</td>
<td>RR 1.65 (1.05, 2.59)</td>
</tr>
<tr>
<td>Racelis 1998</td>
<td>Diet score</td>
<td>Telephone (n=11): improved Control (n=10): improved - “no effect of telephone intervention vs control”</td>
<td>“No significant difference”</td>
</tr>
<tr>
<td>Cummings 1981</td>
<td>Number compliant</td>
<td>All groups (n=93) compliance pre = 86%, post = 90% - “study cannot be used”</td>
<td>Not reported</td>
</tr>
</tbody>
</table>