Grading of Recommendations Assessment, Development, and Evaluation (GRADE) Working Group

www.gradeworkinggroup.org
Is all evidence evaluated equally?

Considerations

- Type of study
- Number of patients
- Quality of research
- Bias & influence
- Strength of effect
- Balance of benefits and risks
- Patient values and preferences
- Role of experience, expertise, consensus

Grading Evidence
A systematic and explicit approach to making judgments about the quality of evidence and the strength of recommendations can help to prevent errors, facilitate critical appraisal of these judgments, and can help to improve communication of this information.
How to Grade Recommendations

**Strong** recommendations
- strong methods
- large precise effect
- few downsides of therapy
- expect non-variant clinician and patient behavior
  - diminished role for clinical expertise
  - focus on implementation & barriers
- focused role of patient values and preferences
  - emphasis on compliance and barriers

**Weak** recommendations
- weak methods
- imprecise estimate
- small effect
- substantial downsides
- expect variability in clinician and patient actions
  - clinical expertise important
  - focus on decision-making and implementation
  - patient values and preferences important
  - focus on determining values and preferences relative to decision

from Holger Schünemann
Grading the Evidence

• Evidence concepts
  – scientific results that approximate truth
  – size, accuracy, precision
  – reliability, reproducibility, appropriateness, bias
  – statistical descriptions
  – trade-offs, limiting factors, cost

• Grade components
  – Quality (Validity)
    • The quality of evidence indicates the extent to which one can be confident that an estimate of effect is correct.
  – Strength (Benefit/Risk)
    • The strength of a recommendation indicates the extent to which one can be confident that adherence to the recommendation will do more good than harm.
The 5 steps in this approach, which follow these judgments, are to make sequential judgments about:

- Which outcomes are critical to a decision
- The quality of evidence across studies for each important outcome
- The overall quality of evidence across these critical outcomes
- The balance between benefits and harms
- The strength of recommendations
Create / produce an evidence summary

- Choose critical outcomes first (define)
  - Typically use an existing systematic review (alternatively can start with other evidence synthesis, or search for original literature, or supplement existing evidence summary with additional evidence about other outcomes)
  - Specify population (subpopulation), & interventions

- Complete an evidence summary
  - GRADEpro facilitates completion of a summary of findings evidence table, with quality grading

- Having included all critical outcomes, it will be possible to judge balance of benefits and risks
**Author(s):** Gunn Vist, Holger Schunemann, Andy Oxman  
**Date:** 21.12.2004  
**Question:** Should Selective Serotonin Reuptake Inhibitors (SSRIs) vs Tricyclic antidepressants be used for the treatment of moderate depression in primary care?  
**Patient or population:** Moderately depressed adult patients  
**Settings:** Primary care  

<table>
<thead>
<tr>
<th>Quality assessment</th>
<th>Summary of findings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No of patients</strong></td>
<td><strong>Effect</strong></td>
</tr>
<tr>
<td><strong>Selective Serotonin Reuptake Inhibitors (SSRIs)</strong></td>
<td><strong>Tricyclic antidepressants</strong></td>
</tr>
<tr>
<td><strong>No Limitations</strong></td>
<td><strong>Relative (95% CI)</strong></td>
</tr>
<tr>
<td>Depression severity (measured with Hamilton Depression Rating Scale after 4 to 12 weeks Range: 0 to 57. Better indicated by: lower scores)</td>
<td></td>
</tr>
<tr>
<td>99 Randomised trials</td>
<td>No limitations</td>
</tr>
<tr>
<td>Transient side effects resulting in discontinuation of treatment (Follow up: 4 to 12 weeks)</td>
<td></td>
</tr>
<tr>
<td>123 Randomised trials</td>
<td>No limitations</td>
</tr>
<tr>
<td>Poisoning fatalities (Follow up: per year of treatment)</td>
<td></td>
</tr>
<tr>
<td>1 Observational studies</td>
<td>Serious limitations (-1)(^2)</td>
</tr>
</tbody>
</table>

Footnotes:  
1. There was uncertainty about the directness of the outcome measure because of the short duration of the trials; it is possible that people at lower risk were more likely to have been given SSRIs and it is uncertain if changing antidepressant would have deterred suicide attempts. RR = 0.02. There is uncertainty about the baseline risk for poisoning fatalities.
Quality of Evidence

The extent to which one can be confident that an estimate of effect or association is correct.

This depends on the:
  – study design
  – study quality (critical appraisal: protection against bias; e.g. concealment of allocation, blinding, follow-up)
  – consistency of results
  – directness of the evidence including:
    • populations (those of interest versus similar; for example, older, sicker or more co-morbidity)
    • interventions (those of interest versus similar; for example, drugs within the same class)
    • comparison (A - C versus A - B & C - B)
    • outcomes (important versus surrogate outcomes)

from Holger Schünemann
Quality of Evidence

- study design *details*
  - study type (e.g. RCT, cohort study, case series)
    (RCT vs observational design *(not* MA, SR, expert)
    - randomization
    - observational study
  - detailed design and execution
    - concealment
    - balance in known prognostic factors
    - intention to treat principle observed
    - blinding
    - completeness of follow-up

from Gordon Guyatt
Quality of Evidence

• consistency of results details
  (similarity of effect across studies)

  – if inconsistency, look for explanation
    • patients, intervention, outcome, methods

  – no clear threshold
    • size of effect, confidence intervals, statistical significance

from Gordon Guyatt
Quality of Evidence

- directness: Patients *details*
  (people, intervention & outcome similar to those of interest)
  - patients meet trials’ eligibility criteria
  - not included, but no reason to question
    - slight age difference, comorbidity, race
  - some question, bottom line applicable
    - valvular atrial fibrillation
  - serious question about biology
    - heart failure trials applicability to aortic stenosis

from Gordon Guyatt
Quality of Evidence

- **directness**: Intervention *details*:
  - similar drugs and doses
  - same class and biology
  - same drugs and doses
  - questionable class and biology

from Gordon Guyatt
Quality of Evidence

• directness: Comparison details:
  – indirect treatment comparisons
    • interested in A versus B
    • have A versus C and B versus C

• Example: alendronate vs risedronate
  – both versus placebo, no head-to-head

from Gordon Guyatt
Quality of Evidence

• directness: Outcomes details:
  – same outcomes
  – similar (duration, quality of life)
  – less breathlessness for role function
  – laboratory exercise capacity for quality of life

from Gordon Guyatt
final considerations:

- magnitude of effect not generally part of quality
  - but very large magnitude can upgrade

- precision not generally part of quality
  - but sparse data can lower quality

- reporting bias
  - high likelihood can lower quality

from Gordon Guyatt
Judgments about Evidence Quality

The quality of evidence indicates the extent to which one can be confident that an estimate of effect is correct.

- **High**: Further research is very unlikely to change our confidence in the estimate of effect.
- **Moderate**: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
- **Low**: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
- **Very Low**: Any estimate of effect is very uncertain.

from Jeff Andrews
Judgments about Evidence Quality

<table>
<thead>
<tr>
<th></th>
<th>Randomized trials</th>
<th>Observational studies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High</strong></td>
<td>No serious flaws in study quality</td>
<td>Extremely strong association and no major threats to validity</td>
</tr>
<tr>
<td><strong>Moderate</strong></td>
<td>Serious flaws in design or execution or quasi-randomized trials</td>
<td>Strong, consistent association and no plausible confounders</td>
</tr>
<tr>
<td><strong>Low</strong></td>
<td>Very serious flaws in design or execution</td>
<td>No serious flaws in study quality, well-done</td>
</tr>
<tr>
<td><strong>Very Low</strong></td>
<td>Very serious flaws and at least one other serious threat to validity</td>
<td>Serious flaws in design and execution</td>
</tr>
</tbody>
</table>

The quality of evidence indicates the extent to which one can be confident that an estimate of effect is correct.

(from Jeff Andrews)
Judgments about Evidence Quality

The quality of evidence indicates the extent to which one can be confident that an estimate of effect is correct.

Moving Up
- very strong association, up 2 levels
- strong, consistent association with no plausible confounders, up 2 levels
- strong association, dose-response can move up 1 level

Moving Down
- study execution, sparse data: serious flaws can lower by 1 level, fatal flaws can lower by 2 levels
- consistency: important inconsistency can lower by 1 level
- directness of evidence: some uncertainty lower by 1 level, major uncertainty lower by 2 levels
- selection bias: strong evidence lower by 1 level

adapted from Gordon Guyatt
Judgments about Evidence Quality

The quality of evidence indicates the extent to which one can be confident that an estimate of effect is correct.

from Jeff Andrews
Strength of Evidence

- What is the magnitude of benefit, and how reliable/precise are these results?
- What is the magnitude of risk & harms & burdens, and how reliable/precise are these results?
- Do the benefits outweigh the risks & harms & burdens? Are there known trade-offs? Are there unknown possible trade-offs?
- Factors that influence strength:
  - Evidence for less serious event than one hopes to prevent
  - Smaller Treatment Effect
  - Imprecise Estimate of Treatment Effect
  - Low Risk of Target Event
  - Higher Risk of Therapy
  - Higher Costs
  - Varying Values
  - Higher Burden of Therapy

from Holger Schünemann
## Strength of Evidence

<table>
<thead>
<tr>
<th>Issue</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evidence for less serious event than one hopes to prevent</td>
<td>Preventing post-phlebitic syndrome with thrombolytic therapy in DVT rather than death from PE.</td>
</tr>
<tr>
<td>Smaller Treatment Effect</td>
<td>Clopidogrel versus aspirin leads to a smaller stroke reduction in TIA (8.7% RRR²) than anticoagulation versus placebo in AF (68% RRR)</td>
</tr>
<tr>
<td>Imprecise Estimate of Treatment Effect</td>
<td>ASA versus placebo in AF has a wider confidence interval than ASA for stroke prevention in patients with TIA.</td>
</tr>
<tr>
<td>Higher Risk of Target Event</td>
<td>Some surgical patients are at very low risk of post-operative DVT and PE while others surgical patients have considerably higher rates of DVT and PE</td>
</tr>
<tr>
<td>Higher Risk of Therapy</td>
<td>ASA and clopidogrel in acute coronary syndromes have a higher risk for bleeding than ASA alone.</td>
</tr>
<tr>
<td>Higher Costs</td>
<td>TPA has much higher cost than streptokinase in acute MI.</td>
</tr>
<tr>
<td>Varying Values</td>
<td>Most young, healthy people will put a high value on prolonging their lives (and thus incur suffering to do so); the elderly and infirm are likely to vary in the value they place on prolonging their lives (and may vary in the suffering they are ready to experience to do so).</td>
</tr>
<tr>
<td>Burden of Therapy</td>
<td>Taking adjusted-dose warfarin is associated with a higher burden than taking aspirin; warfarin requires monitoring the intensity of anticoagulation and a relatively constant dietary vitamin K intake</td>
</tr>
</tbody>
</table>

from Holger Schünemann
When balancing benefits and risks-costs-burdens, the a priori establishment of critical outcomes, and the creation of the evidence profile / summary of findings tables are very useful.
**Author(s):** Gunn Vist, Holger Schunemann, Andy Oxman  
**Date:** 21.12.2004  
**Question:** Should Selective Serotonin Reuptake Inhibitors (SSRIs) vs Tricyclic antidepressants be used for the treatment of moderate depression in primary care?  
**Patient or population:** Moderately depressed adult patients  
**Settings:** Primary care  

<table>
<thead>
<tr>
<th>Quality assessment</th>
<th>Summary of findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of studies</td>
<td>Design</td>
</tr>
<tr>
<td>No patients</td>
<td>Effect</td>
</tr>
</tbody>
</table>

**Depression severity** (measured with Hamilton Depression Rating Scale after 4 to 12 weeks Range: 0 to 57. Better indicated by: lower scores)

| 99 | Randomised trials | No limitations | No important inconsistency | Some uncertainty (-1) | None | 5044 | - | WMD 0.034 (-0.0007 to 0.075) | Moderate | 9 |

**Transient side effects resulting in discontinuation of treatment** (Follow up: 4 to 12 weeks)

| 123 | Randomised trials | No limitations | No important inconsistency | No uncertainty | None | 1948/7032 (27.7%) | 2072/6334 (32.7%) | RR 0.87 (0.80 to 0.95) | 43/1000 (16 to 65) | High | 7 |

**Poisoning fatalities** (Follow up: per year of treatment)

| 1 | Observational studies | Serious limitations (-1) | No important inconsistency | No uncertainty | Very strong association (+2) | 1/100000 (0%) | 58/100000 (0.1%) | RR 0.02 (0.01 to 0.03) | 568/1000000 (0.056% to 0.0574) | Moderate | 8 |

Footnotes:
1. There was uncertainty about the directness of the outcome measure because of the short duration of the trials; it is possible that people at lower risk were more likely to have been given SSRIs and it is uncertain if changing antidepressant would have deterred suicide attempts. RR = 0.02. There is uncertainty about the baseline risk for poisoning fatalities.
Balance between benefits and harms

- **Net benefits:** The intervention does more good than harm.
- **Balance of Trade-offs:** There are important trade-offs between the benefits and harms.
- **Uncertain trade-offs:** It is not clear whether the intervention does more good than harm.
- **No net benefits:** The intervention does not do more good than harm.
- **Net harms:** The intervention does more harm than good.
Judgments about Evidence Strength

The strength of a recommendation indicates the extent to which one can be confident that adherence to the recommendation will do more good than harm.

from Jeff Andrews
Judgments about Evidence Strength

The strength of a recommendation indicates the extent to which one can be confident that adherence to the recommendation will do more good than harm.

from Jeff Andrews
Judgments about Evidence Strength

The strength of a recommendation indicates the extent to which one can be confident that adherence to the recommendation will do more good than harm.

Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group

www.gradeworkinggroup.org

from Jeff Andrews
Judgments about Evidence Strength

The strength of a recommendation indicates the extent to which one can be confident that adherence to the recommendation will do more good than harm.

Net Benefits

Net Harms

Balance of Trade-Offs

Benefits

Absolute Benefit Increase \times Utility

Risks & Harms

Absolute Risk Increase \times Utility (Value Factor)

from Jeff Andrews
Judgments about Evidence **Strength**

The strength of a recommendation indicates the extent to which one can be confident that adherence to the recommendation will do more good than harm.

from Jeff Andrews
Judgments about Recommendations

Judgments about the strength of a recommendation require consideration of:

- all critical outcomes (must be critical, not just important)
- the quality of the evidence
- the balance between benefits and harms
- translation of the evidence into specific circumstances
- the certainty of the baseline risk

- also important to consider costs (resource utilization) prior to making a recommendation

- GRADE suggests using evidence profiles as the basis for making the judgments outlined above
Judgments about Recommendations

- Only outcomes critical to a decision should provide a basis for recommendation
- If information on harm is critical, it should be included even if uncertainty exists
- The lowest quality of evidence for any critical outcome should provide the basis for grading
- However, if evidence favors the same alternative and there is high quality for some but not all of those outcomes, overall quality should still be high
- Weak evidence about implausible putative harms should not lower the overall grade of evidence
<table>
<thead>
<tr>
<th>Determinations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of evidence (GRADE quality assessment with <strong>GRADEpro</strong>)</td>
</tr>
<tr>
<td>Baseline risks of outcomes: benefits, harms, burdens</td>
</tr>
<tr>
<td>Magnitude of relative risks for outcomes: benefits, harms, burdens</td>
</tr>
<tr>
<td>Absolute magnitude of effect for outcomes: benefits, harms, burdens</td>
</tr>
<tr>
<td>Precision of estimates of the effect for outcomes: benefits, harms, burdens</td>
</tr>
<tr>
<td>Relative values of outcomes: benefits, harms, burdens</td>
</tr>
<tr>
<td>Modifying factors that modify effects in specific settings</td>
</tr>
<tr>
<td>Local factors that may affect translating evidence into practice</td>
</tr>
<tr>
<td>Costs</td>
</tr>
</tbody>
</table>

---

**Recommended process for consideration of adjustment to recommendation judgments**

1. **Evidence**
   - Baseline risks of outcomes: benefits, harms, burdens
   - Magnitude of relative risks for outcomes: benefits, harms, burdens
   - Absolute magnitude of effect for outcomes: benefits, harms, burdens
   - Precision of estimates of the effect for outcomes: benefits, harms, burdens
   - Relative values of outcomes: benefits, harms, burdens
   - Modifying factors that modify effects in specific settings
   - Local factors that may affect translating evidence into practice

2. **Action**

---

**Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group**

[www.gradeworkinggroup.org](http://www.gradeworkinggroup.org)
Rationale for downward (weakened) adjustment to recommendation judgments

- Moderate or low quality of evidence
- Uncertainty about baseline risks of outcome
- Balance or important trade-offs between outcomes (inherent value judgment)
- Uncertainty about variation in patients’ relative values of outcomes
- Uncertainty about modifying factors in specific settings
- High Cost / Net Benefits
Judgments about Recommendations

- Do It
- Probably Do It
- No Recommendation
- Probably Don’t Do It
- Don’t Do It
Judgments about Recommendations

- Do It: Strong + For
  - we recommend, should
- Probably Do It: Weak + For
  - we suggest, might, consider
- No Recommendation:
  - (insufficient evidence to support a recommendation for or against)
- Probably Don’t Do It: Weak + Against
  - we suggest, might
- Don’t Do It: Strong + Against
  - we recommend, should
Judgments about Recommendations

Do it = à mettre en pratique

Probably do it = probablement à mettre en pratique

Probably don’t do it = probablement à éviter

Don’t do it = à éviter

From Dominique Broclain
Judgments about Recommendations

- Do it =
- Probably do it =
- Probably don’t do it =
- Don’t do it =

From Jan Brożek, Roman Jaeschke, Wiktoria Leśniak
Judgments about Recommendations

Do it =

• Probably do it =
• Probably don’t do it =
• Don’t do it =

tu es

From Jan Brożek, Roman Jaeschke, Wiktoria Leśniak
Judgments about Recommendations

• **Strong: Do It  or  Don’t Do It**
  – Text: “we recommend” OR “should”
  – Indicating a judgment that a majority of well informed people will make the same choice (high confidence, low uncertainty)
  – Most patients should receive the intervention
    • Could be used as a performance / quality indicator
  – Decision aids not likely to be needed
  – Medical practice is expected to not to vary much

• **Weak: Probably Do It  or  Probably Don’t Do It**
  – Text: “we suggest” OR “might” OR “consider”
  – Indicating a judgment that a majority of well informed people will make the same choice, but a substantial minority will not (significant uncertainty)
  – Decision aids likely to be useful
    • Offering the intervention and helping patients make a decision could be used a quality criterion
  – Medical practice is expected to vary to some degree

• **No Recommendation**
  – Insufficient evidence to support a recommendation; must use clinical expertise and incorporate values and preferences
Judgments about Recommendations

Overall Judgments about Evidence-Based Recommendations

- Do It
- Probably Do It
- No Recommendation
- Probably Don’t Do It
- Don’t Do It

Judgments about Evidence **Strength**

The strength of a recommendation indicates the extent to which one can be confident that adherence to the recommendation will do more good than harm.

Judgments about Evidence **Quality**

The quality of evidence indicates the extent to which one can be confident that an estimate of effect is correct.

Net benefits
Trade-offs
Uncertain trade-offs
No net benefits
Net harm

**High**
**Moderate**
**Low**
**Very Low**

- Strong
- Weak
- None
- Strong
<table>
<thead>
<tr>
<th>Evidence Quality</th>
<th>Net Benefits or Net Harms</th>
<th>Trade-Offs</th>
<th>Uncertain balance or Equal balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Strong Recommendation</td>
<td>Strong Recommendation + Evaluate Values &amp; Preferences</td>
<td>Evaluate Values &amp; Preferences</td>
</tr>
<tr>
<td>Moderate</td>
<td>Weak Recommendation</td>
<td>Weak Recommendation + Evaluate Values &amp; Preferences</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>No Recommendation</td>
<td>No Recommendation</td>
<td></td>
</tr>
<tr>
<td>Very Low</td>
<td></td>
<td></td>
<td>No Recommendation from Jeff Andrews</td>
</tr>
<tr>
<td>Expert Opinion</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Effect of Cost: Example**

The strength of a recommendation indicates the extent to which one can be confident that adherence to the recommendation will do more good than harm.

---

from Jeff Andrews
Profile: New profile for

Profile cover sheet

Question format: Should [Intervention] be used for [health problem]?

Intervention
new profile

Health problem

Question
Should new profile be used for?

Short profile name
New profile for

Author(s)

Date of last minor update: Friday, January 28, 2005

Date of last substantive update: Friday, January 28, 2005

Patients or population

Setting

Systematic review(s)
**Question format**

Should [intervention] vs [comparison] be used for [health problem]?

**Intervention**
- Selective Serotonin Reuptake Inhibitors (SSRIs)

**Comparison**
- Tricyclic antidepressants

**Health problem**
the treatment of moderate depression in primary care

**Question**
Should Selective Serotonin Reuptake Inhibitors (SSRIs) vs Tricyclic antidepressants be used for the treatment of moderate depression in primary care?

**Short profile name**
Selective Serotonin Reuptake Inhibitors (SSRIs) vs Tricyclic antidepressants for the treatment of moderate depression in primary care

**Author(s)**
Gunn Vist, Holger Schunemann, Andy Oxman

**Date of last minor update**
Friday, January 28, 2005

**Date of last substantive update**
Friday, January 28, 2005

**Patients or population**
Moderately depressed adult patients

**Setting**
Primary care

**Systematic review(s)**
Quality Assessment
- Check for continuous outcome

Number of patients
- Intervention:
  - Selective Serotonin Reuptake Inhibitors (SSRIs)
  - Number of patients with outcome: 5044
- Comparison:
  - Tricyclic antidepressants
  - Number of patients with outcome: 4510

Effect
- Effect measure: Weighted mean difference (WMD)
  - WMD: 0.034
  - 95% CI (confidence limits):
    - Low: -0.0007 to 0.075
    - Better indicated by: lower scores
    - Score range: Low: 0 to High: 57

• continuous outcome
Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group

Quality Assessment

Outcome
- Check for continuous outcome

Number of patients

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selective Serotonin Reuptake Inhibitors (SSRIs)</td>
<td>Tricyclic antidepressants</td>
</tr>
</tbody>
</table>

Number of patients with outcome:
- Intervention: 1 / 1000000
- Comparison: 58 / 1000000

Percentage:
- Intervention: (0.0%)  
- Comparison: (0.1%)  

Effect

Relative (95% CI) RR:
- 0.02

95% CI (confidence limits):
- Low: 0.01 to 0.03

Length of follow up:
- per year of treatment

Absolute (95% CI)
- 563 / 1000000

95% CI (confidence limits):
- Low: 562 to 574

• dichotomous outcome
Profile: Selective Serotonin Reuptake Inhibitors (SSRIs) vs Tricyclic antidepressants for the treatment of moderate depression in primary care

Health problem
the treatment of moderate depression in primary care

Question
Should Selective Serotonin Reuptake Inhibitors (SSRIs) vs Tricyclic antidepressants be used for the treatment of moderate depression in primary care?

Short profile name
Selective Serotonin Reuptake Inhibitors (SSRIs) vs Tricyclic antidepressants for the treatment of moderate depression in primary care

Author(s)
Gunn Vist, Holger Schunemann, Andy Oxman

Date of last minor update
Friday, January 28, 2005

Date of last substantive update
Friday, January 28, 2005

Patients or population
Moderately depressed adult patients

Setting
Primary care

Systematic review(s)
### GRADE Evidence Profile

**Author(s):** Gunn Vist, Holger Schunemann, Andy Osman  
**Date:** 1/28/2005  
**Question:** Should Selective Serotonin Reuptake Inhibitors (SSRIs) vs Tricyclic antidepressants be used for the treatment of moderate depression in primary care?  
**Patient or population:** Moderately depressed adult patients  
**Settings:** Primary care  

<table>
<thead>
<tr>
<th>Quality assessment</th>
<th>Summary of findings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No of patients</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Effect</strong></td>
<td></td>
</tr>
</tbody>
</table>

#### Depression severity (measured with Hamilton Depression Rating Scale after 4 to 12 weeks. Range: 0 to 57. Better indicated by lower scores)

<table>
<thead>
<tr>
<th>Design</th>
<th>Quality</th>
<th>Consistency</th>
<th>Directness</th>
<th>Other considerations</th>
<th>No of patients</th>
<th>Effect</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomised trials</td>
<td>No limitations</td>
<td>No important inconsistency</td>
<td>Some uncertainty (-1)¹</td>
<td>None</td>
<td>5044</td>
<td>4510</td>
<td>-</td>
<td>WMD 0.034 (0.0007 to 0.075)</td>
</tr>
</tbody>
</table>

#### Transient side effects resulting in discontinuation of treatment (Follow up: 4 to 12 weeks)

<table>
<thead>
<tr>
<th>Design</th>
<th>Quality</th>
<th>Consistency</th>
<th>Directness</th>
<th>Other considerations</th>
<th>No of patients</th>
<th>Effect</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomised trials</td>
<td>No limitations</td>
<td>No important inconsistency</td>
<td>No uncertainty¹</td>
<td>None</td>
<td>1948/7032 (27.9%)</td>
<td>2072/6334 (22.7%)</td>
<td>RR 0.87 (0.80 to 0.95)</td>
<td>43/1000 (16 to 65)</td>
</tr>
</tbody>
</table>

#### Poisoning fatalities (Follow up: per year of treatment)

<table>
<thead>
<tr>
<th>Design</th>
<th>Quality</th>
<th>Consistency</th>
<th>Directness</th>
<th>Other considerations</th>
<th>No of patients</th>
<th>Effect</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observational studies</td>
<td>Serious limitations (-1)²</td>
<td>No important inconsistency</td>
<td>No uncertainty</td>
<td>Very strong association (12)³</td>
<td>14/10000 (1%)</td>
<td>52/100000 (0.1%)</td>
<td>RR 0.02 (0.01 to 0.03)</td>
<td>588/1000000 (562 to 574)</td>
</tr>
</tbody>
</table>

**Footnotes:**

1. There was uncertainty about the directness of the outcome measure because of the short duration of the trials.
2. It is possible that people at lower risk were more likely to have been given SSRIs and it is uncertain if changing antidepressant would have deterred suicide attempts.
3. RR = 0.02
4. There is uncertainty about the baseline risk for poisoning fatalities.
Other slides to consider

- more about evidence profiles / summary of findings tables
- GRADE publications
- use of GRADE by various organizations
  - list organizations and uses
  - modifications of GRADE
Judgments about Evidence Quality

The quality of evidence indicates the extent to which one can be confident that an estimate of effect is correct.

from Gordon Guyatt, ACCP, UpToDate
Judgments about Recommendations

- **Do It: Strong + For**
  - we recommend, should (1 or ↑↑)

- **Probably Do It: Weak + For**
  - we suggest, might, consider (2 or ↑?)

- **No Recommendation:**
  - (insufficient evidence to support a recommendation for or against)

- **Probably Don’t Do It: Weak + Against**
  - we suggest, might (2 or ↓?)

- **Don’t Do It: Strong + Against**
  - we recommend, should (1 or ↓↓)