

## Evidence Grading Checklist

### The conclusion is supported by GOOD evidence. (A)

<p>1. Evidence</p> <p>The results are from studies of strong research design for answering the practice question, clear methodology and sufficient sample size. Supporting studies might consist of:</p> <p><u>Treatment / Intervention Studies</u></p> <ul style="list-style-type: none"> <li>• good quality systematic review (SR) of randomized controlled trials (RCTs) with consistent findings<sup>i</sup> and a low risk of bias<sup>ii</sup></li> <li>• SR including several trials combined in a single well-done meta-analysis with consistent findings<sup>i</sup></li> <li>• two or more high quality randomized, controlled trials with a low risk of bias<sup>ii</sup>.</li> </ul> <p><u>Etiology / Prognosis Studies</u></p> <ul style="list-style-type: none"> <li>• SR of cohort studies (with homogeneity) or two or more independent well-done prospective cohort studies with consistent results in the absence of evidence to the contrary, where treatment/exposure effects are sufficiently large and consistent</li> </ul> <p>Note: Evidence might also be in a position statement or practice guideline from a national body or organization reporting results of research studies based on the aforementioned types of research</p>	√
2. Consistency <sup>iii</sup> - results are consistent with minor exceptions at most	
3. Clinical impact <sup>iv</sup> - results are clinically important	
4. Generalizability <sup>v</sup> - results are free of any sufficient doubts about generalizability	
5. Applicability <sup>vi</sup> - results are directly applicable to practice setting	

### The conclusion is supported by FAIR evidence. (B)

<p>1. Evidence:</p> <p>The results are from studies of strong design with minor methodological concerns or from studies with weaker designs for answering the practice question, but results have been confirmed in separate studies and are generally consistent. Supporting studies might consist of:</p> <p><u>Treatment / Intervention Studies</u></p> <ul style="list-style-type: none"> <li>• systematic review (SR) of RCTs with heterogeneity although overall the results support the conclusion</li> <li>• a single RCT with low risk of bias<sup>ii</sup></li> <li>• two or more RCTs with a clinically significant conclusion and unclear risk of bias<sup>ii</sup></li> </ul> <p><u>Etiology / Prognosis Studies</u></p> <ul style="list-style-type: none"> <li>• SR of cohort studies (with homogeneity) or two or more well-done prospective cohort studies with consistent findings<sup>i</sup>.</li> <li>• SR of case-control studies (with homogeneity) or several independent case-control studies with similar conclusions</li> </ul> <p>Note: Evidence might also be in a position statement or practice guideline from a national body or organization reporting results of research studies based on the aforementioned types of research</p>	√
2. Consistency <sup>iii</sup> - there is some uncertainty attached to the conclusion because of minor inconsistencies among the results from the studies but inconsistencies can be explained	
3. Clinical impact <sup>iv</sup> - minor doubt about clinical significance of benefits or harms	
4. Generalizability <sup>v</sup> - there is minor doubt about generalizability	
5. Applicability <sup>vi</sup> - generally applicable to practice setting with few exceptions	

**The conclusion is supported by LIMITED evidence or expert opinion. (C)**

<p>1. Evidence</p> <p>The results are from studies of weak design for answering the practice question or there is substantial uncertainty attached to the conclusion because of inconsistencies among the results from different studies. Supporting studies might consist of:</p> <p><u>Treatment / Intervention Studies</u></p> <ul style="list-style-type: none"> <li>• two or more RCTs with inconsistent results or high risk of bias<sup>ii</sup></li> <li>• non-randomized trial or trial that used historical controls</li> <li>• systematic review (SR) of cohort or case-control studies (with homogeneity) or two or more well-done prospective cohort studies with consistent findings<sup>i</sup></li> </ul> <p><u>Etiology / Prognosis Studies</u></p> <ul style="list-style-type: none"> <li>• SR of cohort and case-control studies (with heterogeneity) or two or more studies with some inconsistent results</li> <li>• results from a single cohort study or two or more case-control studies, unconfirmed by other studies</li> <li>• results from a number of high quality cross-sectional studies, well described case reports or case series</li> </ul> <p>Note: Evidence might also be in a consensus report, a position statement or practice guideline from a national body or organization reporting results of research studies based on the aforementioned types of research.</p>	√
<p>2. Consistency<sup>iii</sup> - inconsistencies among the results from different studies leads to substantial uncertainty about conclusions</p>	
<p>3. Clinical impact<sup>iv</sup> – uncertain or moderate</p>	
<p>4. Generalizability<sup>v</sup> - there is substantial uncertainty about the generalizability</p>	
<p>5. Applicability<sup>vi</sup> - likely applicable to practice setting with some exceptions</p>	

**A conclusion is either not possible or extremely limited because evidence is unavailable and/or of poor quality and/or is contradictory. (D)**

<p>1. Evidence:</p> <p>The results are from a single study with major design flaws or from studies with such contradictory results that conclusions can't be drawn. Alternatively, evidence is lacking from either authoritative sources or research involving humans. Supporting studies might consist of:</p> <ul style="list-style-type: none"> <li>• a very poorly designed and executed trial or intervention</li> <li>• evidence from a single case report, case series, case-control study or ecological study unconfirmed by other studies</li> <li>• anecdotal reports</li> <li>• evidence from a small number of similar quality studies that report contradictory results (e.g. two cohort studies that report opposite associations)</li> <li>• research in the <i>in vitro</i>, <i>ex vivo</i> or animal model</li> </ul>	√
<p>2. Consistency<sup>iii</sup> – usually highly inconsistent</p>	
<p>3. Clinical impact<sup>iv</sup> - difficult to assess or minimal</p>	
<p>4. Generalizability<sup>v</sup> – not generalizable or very limited generalizability</p>	
<p>5. Applicability<sup>vi</sup> – not applicable or very limited applicability to the practice setting</p>	

<sup>i</sup> A meta-analysis of RCTs should undergo a statistical analysis of heterogeneity that shows consistency (or homogeneity) between studies.

<sup>ii</sup> Risk of bias is an assessment of the validity of studies included in a review (i.e. the risk that they over- or underestimate the true effect of the intervention). Low risk of bias includes studies that demonstrate adequate sequence generation, allocation

---

concealment, blinding, completeness of outcome data and no other sources of bias. For additional information refer to PEN<sup>®</sup> Writer's Training Module – Appraising the Literature (<http://www.pennutrition.com/WriterGuide.aspx>) and Higgins et al., 2011

<sup>iii</sup> Consistency considers whether findings are consistent across studies, considering the range of study populations and study designs, including the direction and size of the effect or degree of association, and the statistical significance.

<sup>iv</sup> Clinical impact considers the potential benefit of applying the recommendation to a population, including: the relevance of the outcomes to the clinical question, the magnitude of the effect, the length of time to achieve the effect, and the risks versus the benefits.

<sup>v</sup> Generalizability considers how well the population, the intervention and the outcomes in the evidence match the population in the practice question being asked. It considers factors such as gender, age, ethnicity, health status, and how the treatment is delivered.

<sup>vi</sup> Applicability considers whether the evidence is relevant to the practice / health care setting. It considers such factors as access, cost issues etc.

**Note:** The quality of the evidence is a major factor determining the grade; however consideration is given to factors that influence findings, including: consistency, impact, generalizability and applicability. In some cases these factors can supersede the evidence base.

## **Description of Study Designs**

### **Review Articles**

A systematic review is “a critical assessment of existing evidence that addresses a focused clinical question, includes a comprehensive literature search, appraises the quality of studies, and reports results in a systematic manner. If the studies report comparable quantitative data and have a low degree of variation in their findings, a meta-analysis can be performed to derive a summary estimate of effect.” (Ebell et al, 2004).

The evidence cited in the systematic review is what should govern the assignment of the grade. The conclusions generated from a systematic review are only as strong as the research studies included in the review. However, a good quality systematic review should also be well designed and executed. It should describe or include the following:

- search strategy used to locate relevant studies
- study inclusion / exclusion criteria
- an appraisal of the quality and validity of the studies included
- process for data abstraction, synthesis and analysis
- any bias, funding sources or author conflict of interest (authors of the included studies and the systematic review).

A narrative review is a nonsystematic overview of a topic. It generally is not an exhaustive or structured review of the literature, it is more susceptible to bias and does not systematically evaluate the quality of included studies according to any pre-determined criteria. It can be used to identify original studies that can be evaluated and reported as evidence. Generally conclusions from narrative reviews are not reported in the evidence; however in some situations (for example, no recent studies are identified or the compiled studies consist of C- or D-Level evidence), the narrative review can be described in the evidence. In this case, the studies cited should be described and used to assign the evidence grade.

### **Randomized Controlled Trials**

They usually demonstrate whether therapeutic agents are beneficial but can also, less frequently, demonstrate harm. The exposed and unexposed groups should be similar in all respects other than intervention and this balance should be maintained throughout. A high quality randomized controlled trial exhibits the following characteristics: allocation concealed, blinding if possible, intention-to-treat analysis, adequate statistical power, adequate follow-up (>80%).

### **Observational studies**

Observational studies are studies in which investigators do not intervene, but observe the course of events and record changes or differences in one characteristic (e.g. whether they received the exposure of interest such as smoking, exercise or vegetable intake) in relation to changes or differences in other characteristics (e.g. disease development, progression or death).

Observational studies include: cohort studies (prospective or retrospective), case-control studies, cross-sectional studies, case reports and case series.

A cohort study follows a defined group of people (the cohort) over time. Outcomes observed in subsets of the cohort who were exposed to a particular factor are compared to outcomes in those not exposed to a particular factor. A prospective cohort study follows participants into the future; a retrospective cohort study identifies subjects from past records and follows them from the time of those records to a certain point in time. A high quality cohort design exhibits the following characteristics: prospective design, adequate size, adequate spectrum of patients, blinding, a consistent well-defined reference standard, good follow-up, and appropriate adjustment for confounders.

A case-control study compares people with a specific disease or outcome of interest (cases) to people without the disease or outcome (controls) to find associations between the outcome and prior exposure to particular risk factors.

A cross-sectional study measures the distribution of a characteristic in a population or sample at a certain point in time (for example: a survey).

A case report or case study describes observations among a single individual.

A case series study describes observations among a series of individuals usually all subject to the same intervention or exposure, though there is no control group.

### **Expert Opinion**

If there is no critical appraisal or supporting evidence to support statements and conclusions it should not be used as evidence unless it is the only reference you have. In such cases it should be disclosed that the statement is based on unsubstantiated expert opinion.

### **Consensus Reports, Position Statements, Practice Guidelines**

If research studies are cited in a consensus report, position statement or practice guideline from a national or international body or organization, the research studies should govern the grade assignment.

### **References**

Ebell MH, Siwek J, Weiss BD, Woolf SH, Susman J, Ewigman B et al. Strength of recommendation taxonomy (SORT): a patient-centered approach to grading evidence in the medical literature. *J Am Board Fam Pract.* 2004 Jan-Feb;17(1):59-67. Abstract available from: <http://www.ncbi.nlm.nih.gov/pubmed/15014055>

Glossary of Cochrane Collaboration and research terms: <http://www.cochrane.org/glossary>

Greer N, Mosser G, Logan G, Halaas GW. A practical approach to evidence grading. *Jt CommJ Qual Improv.* 2000 Dec;26(12):700-12. Abstract available from: <http://www.ncbi.nlm.nih.gov/pubmed/11143209>

Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD et al.; Cochrane Bias Methods Group; Cochrane Statistical Methods Group The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ.* 2011 Oct 18;343:d5928. Available from: <http://www.bmj.com/content/343/bmj.d5928.long>

NHMRC: NHMRC levels of evidence and grades for recommendations for developers of guidelines. Canberra, ACT: National Health and Medical Research Council, Commonwealth of Australia; December, 2009. Available from: [http://www.nhmrc.gov.au/\\_files\\_nhmrc/file/guidelines/developers/nhmrc\\_levels\\_grades\\_evidence\\_120423.pdf](http://www.nhmrc.gov.au/_files_nhmrc/file/guidelines/developers/nhmrc_levels_grades_evidence_120423.pdf)