EBSJ Evidence Assessment: Definition of Risk of Bias (RoB)

**Articles on treatment**

| Risk of bias description | Studies of therapy | Study design | Criteria |
|--------------------------|--------------------|-------------|----------|
| Low risk                 | Study adheres to commonly held tenets of high-quality design, execution, and avoidance of bias | Good quality RCT | • Random sequence generation  
• Allocation concealment  
• Intent-to-treat analysis  
• Blind or independent assessment for author’s primary important outcomes  
• Co-interventions applied equally  
• F/U rate of ≥ 80%  
• Patients followed long enough for outcomes to occur  
• Accounting for other prognostic factors  
• Objective and unbiased outcome measure used |
| Moderately low risk      | Study has potential for some bias; study does not meet all criteria for a good quality RCT, but deficiencies not likely to invalidate results or introduce significant bias | Moderate or poor quality RCT | • Violation of one or two criteria for good quality RCT |
| Moderately high risk     | Study has significant flaws in design and/or execution that increase potential for bias that may invalidate study results | Poor quality RCT | • Violation of three or more of the criteria for a good quality RCT |
| High risk                | Study has significant potential for bias; lack of comparison group precludes direct assessment of important outcomes | Case series | • Any case series design |

**Articles on prognosis or risk**

| Risk of bias description | Studies of prognosis | Study design | Criteria |
|--------------------------|----------------------|-------------|----------|
| Low risk                 | Study adheres to commonly held tenets of high-quality design, execution, and avoidance of bias | Good quality cohort | • Prospective design  
• Patients at similar point in the course of their disease or treatment  
• F/U rate of ≥ 80%  
• Patients followed long enough for outcomes to occur  
• Accounting for other prognostic factors  
• Objective and unbiased outcome measure used |
| Moderately low risk      | Study has potential for some bias; study does not meet all criteria for good quality cohort, but deficiencies not likely to invalidate results or introduce significant bias | Moderate quality cohort | • Prospective design, with violation of one of the other criteria for good quality cohort study  
• Retrospective design, meeting all the rest of the criteria in good quality cohort |
| Moderately high risk     | Study has flaws in design and/or execution that increase potential for bias that may invalidate study results | Poor quality cohort  
• Good-quality case-control or cross-sectional study | • Prospective design with violation of two or more criteria for good quality cohort, or  
• Retrospective design with violation of one or more criteria for good quality cohort  
• A good case-control study  
• A good cross-sectional study  
• Any case series design |
| High risk                | Study has significant potential for bias; does not include design features geared toward minimizing bias and/or does not have a comparison group | Poor quality case-control or cross-sectional Case series | • Other than a good case-control study  
• Other than a good cross-sectional study  
• Any case series design |

**Determination of Overall Strength (Quality) of Evidence (SOE)**

After individual article evaluation, the overall quality of the body of evidence with respect to each outcome is likely to include concepts outlined by the Grades of Recommendation Assessment, Development, and Evaluation (GRADE) Working Group and recommendations made by the Agency for Healthcare Research and Quality (AHRQ). Qualitative analysis is performed considering the AHRQ required and additional domains. The table below provides an outline of the methods used to determine the final SOE.

The following four possible levels and their definition will be reported:

- **High**: High confidence that the evidence reflects the true effect. Further research is very unlikely to change our confidence in the estimate of effect.
- **Moderate**: Moderate confidence that the evidence reflects the true effect. Further research may change our confidence in the estimate of effect and may change the estimate.
- **Low**: Low confidence that the evidence reflects the true effect. Further research is likely to change the confidence in the estimate of effect and likely to change the estimate.
- **Insufficient**: Evidence either is unavailable or does not permit a conclusion.

All AHRQ “required” and “additional” domains are assessed. Only those that influence the baseline grade are listed in table. Baseline strength: Risk of bias (including control of confounding) is accounted for in the individual article evaluations. High = majority of articles RCTs; Low = majority of articles cohort (observational) studies.

**Definitions of the Different Levels of Evidence for Reliability Studies**

Level of evidence ratings for Cochrane reviews and other systematic reviews are assigned a baseline score of High if RCTs were used. Low if observational studies were used. The rating can be upgraded or downgraded based on adherence to the core criteria for methods, quality, and quantitative analyses for systematic reviews (there is a reference/evaluation table for this).

**Level**

1. Good quality study  
2. Moderate quality  
3. Poor quality study  
4. Very poor quality study

**Study type**

- • Broad spectrum of patients with the expected condition  
- • Adequate description of methods for replication  
- • Blinded performance of tests, measurements, or interpretation  
- • Second test/interpretation performed independently of the first

**Criteria**

| 1 | Good quality study |
| 2 | Moderate quality | • Violation of any one of the criteria for a good quality study |
| 3 | Poor quality study | • Violation of any two of the criteria |
| 4 | Very poor quality study | • Violation of all three of the criteria |