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

**Articles on treatment**

| Risk of bias description | Studies of therapy | Criteria |
|--------------------------|--------------------|---------|
| Low risk                 | Study adheres to commonly held tenets of high-quality design, execution, and avoidance of bias | Moderate or poor 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%+; Consistency unknown. |
|                          |                    | 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%+; Consistency unknown. |
| Modestly low risk        | Study has potential for some bias, study does not meet all criteria for a good quality RCT, but deficiencies are not likely to invalidate results or introduce significant bias | Moderate or poor 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%+; Consistency unknown. |
| 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 | Criteria |
|--------------------------|----------------------|---------|
| Low risk                 | Study adheres to commonly held tenets of high-quality design, execution, and avoidance of bias | Good quality cohort | Moderate quality cohort |
|                          |                      | Moderate quality cohort |
|                          |                      | Poor quality cohort |
|                          |                      | Case-control or cross-sectional study |
| Moderate low risk        | Study has potential for some bias; does not meet all criteria for good quality cohort, but deficiencies are not likely to invalidate results or introduce significant bias | Moderate quality cohort |
| Moderate high risk       | Study does not have a comparison group, does not fit within a defined population, or is not adequately powered | Poor quality cohort |
|                          |                      | Good quality case-control or cross-sectional study |
| High risk                | Study has significant potential for bias; does not include features geared toward minimizing bias and/or does not have a comparison group | Other than a good case-control study |
|                          |                      | Other than a good cross-sectional study |
|                          |                      | Any case series design |

\[a\]Outcome assessment is independent of health care personnel, investigator, or patient judgment.

\[b\]Authors must provide a description of robust baseline characteristics, and control for those that are equally distributed between treatment groups. RCTs get credit if there is a similar distribution of baseline characteristics between groups but must also control for confounding if distribution is not similar.

\[c\]Reliable data are data such as mortality or reoperation.

**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 exceed the RoB endpoints 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.

**Strength of Evidence for Existing Systematic Reviews**

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, qualitative, and quantitative analyses for systematic reviews (there is a reference/evaluation table for this).

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

| Level | Study type | Criteria |
|-------|------------|----------|
| 1     | Good quality study | Broad spectrum of persons with the expected condition |
| 2     | Moderate quality | Adequate description of methods for replication |
| 3     | Poor quality study | Adequate description of methods for replication |
| 4     | Very poor quality study | Adequate description of methods for replication |

\[a\]Required domains: risk of bias, consistency, directness, precision. Plausible confounding that would decrease observed effect is accounted for in our baseline risk of bias assessment through individual article evaluation. Additional domains: dose-response, strength of association, publication bias.

\[b\]Single study = "consistency unknown."