Definition of the Different Classes of Evidence (CoE)

**Articles on treatment**

| Class | Risk of bias | Study design | Criteria |
|-------|-------------|--------------|----------|
| I     | Low risk    | Good quality RCT | • Random sequence generation  
• Allocation concealment  
• Intention-to-treat analysis  
• Blind or independent assessment for important outcomes  
• Co-interventions applied equally  
• F/U rate of 80%  
• Adequate sample size |
| II    | Modestly low risk | Moderate or poor quality RCT | • Violation of one of the criteria for good quality RCT  
• Blinding or assessment in a prospective study, or use of reliable data  
• F/U rate of 80%  
• Adequate sample size  
• Controlling for possible confounding |
| III   | Moderately high risk | Moderate or poor quality cohort | • Violation of any of the criteria for good quality cohort  
• Any case-control design |
| IV    | High risk   | Case-control | • Any case series design |

**Articles on prognosis or risk**

| Class | Risk of bias | Study design | Criteria |
|-------|-------------|--------------|----------|
| I     | Low risk    | 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 |
| II    | Modestly low risk | 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 class I |
| III   | Modestly high risk | Poor quality cohort | • Prospective design with violation of 2 or more criteria for good quality cohort study, or  
• Retrospective design with violation of 1 or more criteria for good quality cohort study  
• A good case-control study  
• Other levels of evidence study |
| IV    | High risk   | Poor quality case-control or cross-sectional study | • Other than a good case-control study  
• Other than a good case-control study |

*Outcome assessment is independent of healthcare personnel judgment. Reliable data are data such as mortality or re-operation.*

**Explanation of Strong of Evidence for Existing Systematic Reviews**

Level of evidence ratings for Cochrane reviews and other systematic reviews are assigned a baseline score of high (I/II) or low (III/IV) quality cohort studies. This rating may be upgraded or downgraded based on adherence to the core criteria for methods, qualitative, and quantitative analyses for systematic reviews (there is a reference table for this). The following four possible levels and their definitions are 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 both 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 the table. Baseline strength: Risk of bias (including control of confounding) is accounted for in the individual article evaluations. High = majority of articles level II/III; low = majority of articles level III/IV.

**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  
• Adequate description of methods for replication  
• Blinded performance of tests, measurements or interpretation  
• Second test/interpretation performed independently of the first |
| 2     | Moderate quality | • Violation of any one of the criteria for a good quality study |
| 3     | Poor quality study | • Violation of any of the criteria |
| 4     | Very poor quality study | • Violation of all three of the criteria |

*Required domains: risk of bias, consistency, directness, precision. Placeable 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.*

*Authors must provide a description of robust baseline characteristics, and control for those that are unevenly distributed between treatment groups.*

*Additional domains: risk of bias, consistency, directness, precision. Placeable 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.*