Supplemental figure: Flow chart of decision-making process for levels of evidence, based on study design, study quality and study size.

Consistent positive results (66.6% of relevant studies reporting significant positive results) are needed to achieve strong, moderate or limited levels of evidence.

**Consistent results?**
- Consistent positive results?
  - Yes → Strong evidence
  - No → Moderate evidence

**Level of evidence**
- Consistently no effect?
  - Yes → Limited evidence
  - No → Inconclusive

**Any large high quality RCTs?**
- Yes → ≥2
  - Consistent positive results?
    - Yes → Strong evidence
    - No → No evidence
  - Consistently no effect?
    - Yes → No evidence
    - No → Inconclusive
  - No consistent results?
    - Yes → Inconclusive
    - No → Inconclusive

**Any low quality RCTs (or small high quality RCTs) or large high quality CTs?**
- Yes → ≥1
  - Consistent positive results?
    - Yes → Moderate evidence
    - No → No evidence
  - Consistently no effect?
    - Yes → No evidence
    - No → Inconclusive
  - No consistent results?
    - Yes → Inconclusive
    - No → Inconclusive

**Any low quality CTs (or small high quality CTs)?**
- Yes → ≥1
  - Consistent positive results?
    - Yes → Limited evidence
    - No → No evidence
  - Consistently no effect?
    - Yes → No evidence
    - No → Inconclusive
  - No consistent results?
    - Yes → Inconclusive
    - No → Inconclusive

**Any small high quality CTs or low quality CTs?**
- Yes → ≥1
  - Consistent positive results?
    - Yes → Limited evidence
    - No → No evidence
  - Consistently no effect?
    - Yes → No evidence
    - No → Inconclusive
  - No consistent results?
    - Yes → Inconclusive
    - No → No

**NOTE:** studies including ≤ 250 participants or not providing sample size justifying a smaller sample size are considered ‘small’, studies including >250 participants are considered ‘large.’