Understanding GRADE (Grading of Recommendations Assessment, Development, and Evaluation)

The following concepts are utilized when applying GRADE to a literature analysis that answers a clinical question. The quality of the evidence (High, Medium, Low, or Very Low) reflects the confidence in the estimate of effect for an outcome to support the recommendation. When the evidence is only from randomized control trial (RCT) the GRADE starts as High quality evidence.

The RCTs are read and assessed upon the criteria below.
- Risk of Bias, looks specifically for:
    - Randomization
    - Concealment of allocation
    - Blinding (participants, providers, outcome assessor)
    - How subjects who did not complete the study were handled
    - Selective reporting
- Inconsistency (heterogeneity) is present when:
    - Variation of point estimates across the included studies is reported?
    - The confidence intervals of the included studies do not overlap, or have minimal overlap
    - The p value of the statistical tests for heterogeneity (Chi^2) is low (p< 0.05)
    - The I^2 statistic is roughly interpreted as (note the overlap in the ranges below; I^2 is not a precise measure):
        - 0%-40% the heterogeneity might not be a factor
        - 30-60% may indicate moderate heterogeneity
        - 50-90% may indicate substantial heterogeneity
        - 75-100% indicates considerable heterogeneity
- Indirectness is present when:
    - Studies with a direct comparison are not available, and the included studies are combined to indirectly look at the comparison of interest. Example: The Migraine in the Emergency Department Clinical Practice Guideline compares dexamethasone to magnesium sulfate IV for the outcome pain relief at 2 hours. Studies have been identified that compare (a) dexamethasone to magnesium sulfate (IV), and (b) prochlorperazine to magnesium sulfate (IV). If they are used to compare dexamethasone to prochlorperazine the evidence is indirect (see Figure A1).
    - Studies predominantly include (a) participants, (b) methods- such as doses, (c) outcomes that are different from the question asked by the guideline panel. In pediatrics, this is most often seen when adult population studies are applied to children.
- Imprecision is present when:
    - The included studies have less than ~ 400 subjects cumulatively (this is a rule of thumb)
    - The included studies have a small number of events. For example, for the outcome mortality, if there are only 2 deaths across all studies, the evidence will be downgraded for imprecision.
    - The 95% confidence interval of the cumulative effect includes
        - the line of no effect and the confidence intervals of the included study are wide
- Publication Bias is evident when:
    - Literature is systematically over- or underestimate of the treatment effect due to which studies are selected for publication
        - Studies with small sample sizes are less likely to be published
        - Negative studies are less likely to be published
        - All the studies have strong links to commercial enterprises that might benefit from the results
    - Publication bias is assessed by evaluating funnel plots

EXAMPLE

For each criterion, if present, the quality of evidence is reduced by one. For example, four studies have been located for the comparison dexamethasone versus magnesium sulfate (IV) to treat refractory migraine in the ED. The outcome is Hospitalization. The studies start as High quality for grading summary (See Table A1).

The studies have:

- Risk of Bias- the study is methodologically strong- Low risk of bias- not downgraded
- Inconsistency- confidence intervals overlap and the I^2 is 0%- Low risk of inconsistency- not downgraded (See Figure A2)
- Indirectness- no direct comparison, extrapolated from comparison to another drug- High risk of indirectness- downgraded to Moderate
- Imprecision- there are four studies, the cumulative number of subjects is 299, and the number of events (number of subjects hospitalized is 11). Imprecision is high because (a) the effect size crosses the line of no effect, and (b) the number of subjects and the number hospitalized are low, therefore the evidence has High risk of imprecision- downgraded to Low (See Figure A2.)
- Publication Bias- unable to assess, only four small studies are included- Stays at Low

For this outcome, using the GRADE criteria, the quality of evidence for this comparison is Low.

Table A1. Grade Summary for the Example

Quality of Evidence for this Outcome	GRADE
High	Starting point
Moderate	Downgrade to Moderate due to Indirectness
Low	Downgrade to Low due to Imprecision
Very Low	

Note: this evidence was not downgraded for risk of bias, inconsistency or publication bias.
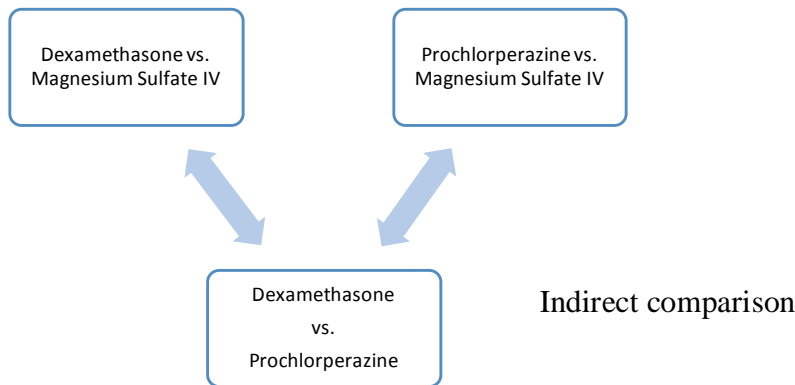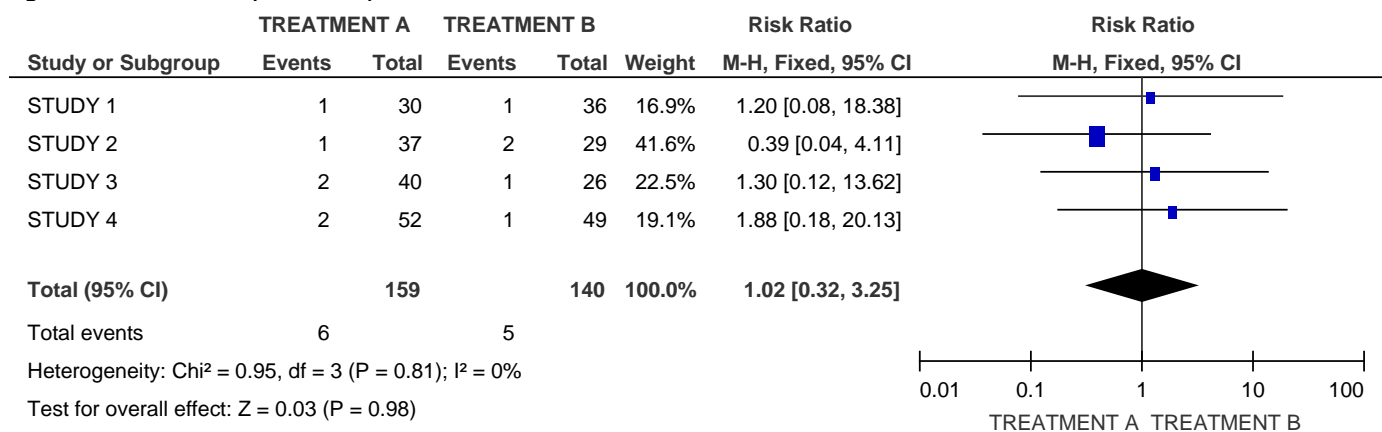
Figure A1. Example of Indirectness

Figure A2. Meta-analysis example for Grade

Reference: Schünemann H, Santesso, N. (2016). How to GRADE the Evidence.   Retrieved from http://training.cochrane.org/