Delfini Pearls
Quickly Evaluating Reliability & Usefulness of Efficacy Evidence for Therapies

I. Should I bother with this article? Do not automatically trust authors or sources. Details must be evaluated.

If the Results are Reliable, Are They Usable? ⬇	**Pertinent Study Issues For Determining Reliability & Usability** ⬇
1. Will I change my practice? Will my patients benefit? 2. Meaningful benefit = size of benefit + areas of clinical significance to patients (morbidity, mortality, symptom relief, physical and emotional functioning and health-related quality of life) 3. Applicability to my patient (see inclusions, exclusions and baseline characteristics) 4. Look at the boundaries of the confidence intervals (CIs) of valid studies to compare to your requirements for meeting clinical significance. See * below. **Then** ➔	5. Was choice involved in determining who got the therapy? If yes, this is an observational study, and there is a high risk of misleading results, with rare exception (e.g., all-or-none results). 6. Is the comparison fair? 7. Question, endpoints and analysis groups should be determined in advance, otherwise there is a high risk of chance results. 8. Do you agree with how they defined outcomes such as success/failure, improvement/no improvement, etc.? 9. If this study is for a new agent, safety may be unknown.

II. How do I know the results are likely to be true?

10. What could possibly explain the results other than the intervention? Could **bias**, **confounding** or **chance** explain or affect the outcomes? Or can I rule them out and comfortably presume results are true (**cause and effect**)?
11. The burden of proof is on the intervention: Could anything advantage the intervention?
12. Review the 4 phases of a study for bias: **selection** phase + **performance** phase + **follow-up** phase + **assessment** phase. Always use a critical appraisal checklist. See the **Delfini Study & Usability Tool** (short checklist) for validity considerations.
    - Bias tends to favor the intervention. Biases in areas such as randomization, concealment of allocation to study groups, blinding or using models can inflate benefits by up to a relative 50 percent.
    - Are there **any differences** between groups except for what is being studied? If yes, this is an automatic bias.
    - Look for other biases. For example, frequently information on co-interventions or adherence will be lacking. This could result in misleading results due to confounding. Were methods used to measure the outcomes appropriate, etc.?
    - How much **data** are **missing**? (e.g., discontinuations, etc.)? Even non-differential loss can mean differences in prognostic variables. Primary analysis should be Intention-to-Treat (ITT), with people being analyzed in their assigned groups plus using reasonable choices to "make up" entries for missing data. Most analyses involve modeling and modeling requires unverifiable assumptions.

Best Experiments for Interventions: Prevention & Therapies ^

a. Questions in advance (a priori) and questions that matter to patients (clinical significance)
b. Right people, enough people and low number of missing data points (possible to solve via analysis)
c. Comparison group (concurrent controls) and good comparator (eg, placebo)
d. Randomization to assign study groups
e. Blinding successful (assignment to groups, all people, intervention, all working with data)
f. Reasonable design, execution, measurement and analysis methods (including composite endpoints)
g. Sameness except for intervention
h. No bias or confounding
i. Tests for probability of coincidental findings (chance)
j. Results large enough to be useful or, if no difference between groups, assurance results are conclusive (review confidence intervals)
k. Safety considered

^ Special issues: cross-over design, screening, diagnostic testing, equivalence and non-inferiority trials

III. For valid studies, how do I know the results are likely to be useful?

13. *For valid studies, consider what *you judge* to be a reasonable range for **clinical significance**. For **statistically significant findings**, is the confidence interval wholly within *your* bounds for clinical significance? For **non-significant findings**, is the confidence interval wholly beneath your limit for clinical significance? A yes to these two questions means likely conclusive findings for valid studies. No, means findings are inconclusive.
14. **Non-significant findings** may mean there truly is no difference or they didn't have enough people to show a difference.
15. Equivalence does not mean two drugs are actually equivalent. There are special issues for evaluating "equivalence" and "non-inferiority" available from Delfini.

IV. Suggestions for quality information sources are available from Delfini. Most all sources vary in quality and should be critically appraised. Delfini has tools available for appraising primary and secondary sources for interventions.