How to Interpret a Clinical Trial Result

Edward V. Loftus, Jr., M.D.
Professor of Medicine
Division of Gastroenterology and Hepatology
Mayo Clinic College of Medicine
Rochester MN
Overview

• Are results valid?
  — Randomization
  — Stratification
  — Concealed Allocation
  — Blinding

• What are the results?
  — Primary vs. secondary endpoint
  — Pre-specified vs. post-hoc analysis
  — How large was treatment effect?
    • Absolute benefit, relative benefit, number needed to treat
Overview

• How can I apply the results to patient care?
  — Type of study patients
  — Clinically important outcome?
  — Are benefits worth the harms and costs?

• Disclosures/Conflicts of Interest
IBD Therapy Trial Design

• Induction
  — 4 to 12 weeks
  — Balance between onset of action and placebo response

• Maintenance
  — 26 to >52 weeks
  — Randomize from the start vs. open label drug then randomize responders (e.g., PRECiSE 1/ACCENT 2/ACT 1-2 vs. PRECiSE 2/ACCENT 1/CHARM)
IBD Therapy Endpoints

• Crohn’s Disease Activity Index (CDAI)
  — 7-day symptom diary
  — Need weight and Hgb, plus physical exam for abdom mass
  — Increasing recognition this may correlate poorly with objective markers

• Ulcerative colitis – Mayo score or Sutherland score
  — Rectal bleeding, stool frequency, endoscopic severity, physician global assessment
Evidence-Based Medicine Approach to Articles of Therapy - Validity

- Randomization
- Concealed Allocation
- Double Blinding
- Complete Follow-Up of Patients
- Intention to Treat Analysis
Randomization

- Randomization: Patient has predetermined chance to get new treatment or control treatment
- Why use randomization? Many known factors (e.g., age, co-morbid illness) and unknown factors may influence outcome
Randomization

- Randomization attempts to eliminate bias by distributing these factors evenly between new treatment and control groups.
- For example, randomization should lead to equal distribution of patients with severe flares of Crohn’s disease in maintenance and placebo groups.
Concealed Allocation

- When enrolling a patient into a trial, the researcher obtaining informed consent does not know if the next patient will get new treatment or control (e.g., opaque envelope).
- Concealed allocation is an extension of randomization.
Concealed Allocation: Example

- If the investigator or study coordinator knows which treatment the next study patient will receive, that might influence their decision to offer the study to a particular patient.
- Example, the next potentially eligible patient, with a history of refractory disease is interviewed by the nurse.
- She is concerned that he will relapse if he doesn’t continue to receive a particular medication.
Concealed Allocation: Example

• Without concealed allocation, the nurse knows the patient will get placebo if he enters the trial.

• The nurse may therefore subconsciously try to convince the patient not to enroll in the trial.

• Ultimate outcome: Patients with refractory disease will not be evenly divided between the two groups.
Bias in Clinical Trials

Double Blinding

- Patients and healthcare providers (and/or the assessors of the study outcome) do not know if a patient is getting new treatment or control.

- Double blinding prevents bias in administration of concurrent therapy or the interpretation of outcome.
Double Blinding

• Example: If a patient knows he/she is receiving the study drug, this knowledge may influence his/her subjective assessment of efficacy measurements and side effects.

• This is particularly relevant for an endpoint such as the CDAI, which is primarily influenced by symptoms.
Key Issues in Studies of Therapy

• Epidemiologic studies have proven that randomization, concealed allocation, and double blinding prevent inflated estimates of treatment benefit.

Prognostic Factors at Baseline

• Were patients in the groups similar with respect to known prognostic factors?

• A good Table 1 will list baseline factors for all study groups.

• If these differences occur, they can sometimes be adjusted for in the analysis.
Complete Follow-Up of Patients

• When a patient is lost to follow-up, two possibilities must be considered:
  — Patient is healthy and doesn’t need to follow-up
  — Patient is ill and too sick to follow-up

• If many patients are lost to follow-up, then results of trial may not be accurate.
Complete Follow-Up of Patients

• How do you determine if too many patients have been lost to follow-up?
  — No easy ‘rules of thumb’, depends on rate of study outcome
  — The fewer the events the more serious incomplete follow-up is

• Re-calculate results assuming that patients lost to follow-up in new treatment group had bad outcome and patients lost to follow-up in control group had good outcome.
Intention to Treat (ITT) Analysis

• Hypothetical Example: Comparison of CABG vs. medical therapy for the treatment of patients with angina with death as outcome

• If a patient is randomized to get CABG, but the patient dies before the procedure, should the patient be included in the final data analysis? If so, in which group?
Intention to Treat Analysis vs. Per Protocol Analysis

- **Intention to Treat Analysis:**
  - All randomized patients with known outcomes are included in final data analysis, regardless of whether they complete trial.
  - Preserves the value of randomization.

- **Per Protocol Analysis:**
  - Only patients who complete the trial according to protocol are analyzed.
  - Often used when examining adverse events associated with study drug.
**Sample Size Calculation**

- Most good RCTs will have a sample size calculation in the Methods.
- Make assumptions about treatment outcomes and loss of follow-up in each study group, and back-calculate to show you have 80% power to demonstrate a statistically significant difference.
- This becomes very relevant for a negative study: was it truly negative or just underpowered?
Sample Size Calculation Part 2

• Superiority vs. non-inferiority (formerly known as ‘equivalence’ trial)
• Need MUCH larger samples to prove 2 treatments are equivalent than to prove one is superior over the other
Results: Keep Your Eye on the Ball!

• ‘The ball’ is the pre-specified primary endpoint.
• Trial results should be stated first and foremost with respect to the primary endpoint.
• A trial that misses its primary endpoint but makes various secondary/post-hoc endpoints is fundamentally a negative trial.
Results: Secondary Endpoints

- Secondary endpoints can be interesting but should be considered ‘hypothesis generating’ and not ‘confirmatory’
  - Example: ENACT-1 (natalizumab induction) missed primary endpoint but post-hoc analysis suggested those with elevated CRP were more likely to respond. This led to ENCORE (natalizumab induction in high-CRP patients) which confirmed the finding.
Determining the Magnitude of Treatment Effect

- Absolute Benefit Increase (ABI)
- Relative Benefit Increase (RBI)
- Number Needed to Treat (NNT)
Absolute Benefit Increase (ABI)

• Actual increase in good outcomes between treatment group patients & control group patients
• (% good outcome: treatment group) - (% good outcome: control group) = ABI
• 80% - 49% = 31%
• Thus, an extra 31% of patients with a history of erosive esophagitis will remain free of esophagitis if they take omeprazole instead of ranitidine.
Relative Benefit Increase (RBI)

- Increased chance of good outcome in treatment group patients compared to chance of good outcome in control group patients

- \[ \frac{(% \text{good outcome: tx group}) - (% \text{good outcome: control group})}{(% \text{good outcome: control group})} \]

- \[ \frac{(80\% - 49\%)}{49\%} = 63\% = \text{RBI} \]

- In other words, a patient with a history of erosive esophagitis who receives omeprazole is 63% more likely to remain free of esophagitis compared to a similar patient who receives ranitidine.
Number Needed to Treat (NNT)

- Number of patients that need to receive treatment rather than control for one additional good outcome to occur over a specified period of time

\[
NNT = \frac{1}{\text{ABI}} = \frac{1}{0.31} = 3
\]

Thus, for every three patients treated with omeprazole instead of ranitidine, one additional patient will remain in remission from erosive esophagitis.
Determining the Precision of Treatment Effect

- p values
- 95% confidence intervals
Determining the Precision of Treatment Effect

- **MUCOSA Trial:**
  RCT of rheumatoid arthritis patients using NSAIDs. Patients given co-therapy with misoprostol or placebo and followed for serious GI complications.

- **Serious GI Complications:**
  - placebo: 42/4439 or 0.95%
  - misoprostol: 25/4404 or 0.57%
$p$ Values: Did the Difference Between Treatment & Placebo Occur Due to Chance?

MUCOSA Trial: RRR = 40%; $p = 0.05$

- $p = 0.05$: 5% probability that observed difference between treatment & placebo occurred due to chance

- $p = 0.05$: 95% probability that observed difference between treatment & placebo occurred because there is a true difference
95% Confidence Intervals: Precision of Results

- MUCOSA Trial: RRR = 40%; 95% CI: 1-64%

- If the trial was repeated 100X, in 95/100 trials, the RRR would fall between 1% & 64%

- Simpler definition: 95% probability that true RRR lies between 1% & 64% (i.e., 95% CI)
Statistical Significance vs. Clinical Significance

- If trial demonstrates 40% RRR & 0.4% ARR with \( p \) value = 0.04, then treatment is statistically better than placebo.

- Is this difference between treatment & placebo clinically significant? (Hint: calculate the NNT)
Applying Results to a Specific Patient

- Is the patient population in the study similar to your patient?
- Do the benefits of treatment with the study drug outweigh the side effects in your patient?
Disclosures/Conflict of Interest

• Full disclosure of potential conflicts of interest allows the fully informed reader to make decision about validity of results.

• Most journals now require disclosure (past 1-2 years) of research support, consulting fees, speakers’ bureau, CME events indirectly sponsored, stock ownership

• IMPORTANT: sponsorship by a company does NOT mean the study results are more suspect
  — Companies are often the only entities with pockets deep enough to perform an adequately powered trial
Interpreting RCTs: Summary

- Make sure the trial meets validity criteria: randomization, concealed allocation, double-blinding, ITT analysis, adequate follow-up.
- Make sure the trial is adequately powered, especially if it’s a negative trial: look for sample size calculation.
- Ensure results are stated in terms of pre-specified primary endpoint.
Interpreting RCTs: Summary

- Think of results in terms of absolute benefit increase, NNT, and relative benefit increase.
- 95% CI preferable to p-values.
- Do the study patients reflect your patients?
- Be cognizant of disclosures/COI.