**University of Utah Internal Medicine Journal Club Template:**

Updated: Brian Locke 2/25/2020

**RANDOMIZED CONTROLLED TRIAL FACILITATOR’S GUIDE**

Randomized Controlled Trial (Assigned Resident):

* Article Title
* Study question and design: (Is this question important? Run-in period? Parallel Groups, Factorial, Cross-over, or other design?)
* Patients included: (where and how were patients enrolled? How did the trialists determine their sample size? What were the inclusion/exclusion criteria? Does this represent the population you’re interested in? Flow diagram and table 1 may be helpful)
* Intervention: (What was the experimental intervention? Aside from the intervention, were patients treated equally - follow-up schedule, permitted additional treatments etc? Does this reflect current practice? Is the intervention realistic in our setting?)
* Outcomes: (Focus on primary outcome. Is this an outcome important to patients, or a surrogate? Is it a composite? Appropriate duration of follow-up? Was there blinding, and of who? Were patients analyzed by the treatment they were assigned to, or the treatment they received? Were the secondary outcomes prespecified or post-hoc?)
* Results: How large was the treatment effect, and what was the precision in the estimate (e.g. confidence interval, P-value)? Was there significant loss to follow-up or crossover between groups? Were the secondary analyses prespecified, or post-hoc? What were the primary adverse events, and how common were they in each group?
* Critique: Are there threats to the internal validity (such as bias or chance) or external validity, aka generalizability? Relevant conflicts of interests?
* Can I apply the results to my patient? How?

Common discussion points for critiques of RCTs:

General points about RCTS:

*Advantages:*

* unbiased distribution of confounders;
* blinding more likely;
* randomization facilitates statistical analysis (high quality data, comparator group present)

.

*Disadvantages:*

* expensive: time and money;
* volunteer bias;
* ethically problematic at times (such as if no equipoise between groups)

**Study question and design (Facilitator):**

-was this RCT prespecified e.g. clincicaltrials.gov. Why is this important? (all should be, to avoid biasing publications toward positive trial results). Additionally, this prespecifies the study hypotheses and design decisions.

-How was randomization done? The Results section should say how many patients were randomized (eg Baseline Characteristics table) = best is centralized, with concealed allocation. Was randomization concealed? (if not, there is a chance investigators could enroll sicker or less sick patients to one of the groups)

**Patients included (Facilitator):**

-do all studies need multicenter confirmation before results are applicable? Can argue so, especially if local expertise / demographics / environment could limit generalizability

-What assumptions did they make about the treatment effect size in order to calculate sample? Usually, they will use a prior trial or cohort to estimate the effect size. Smaller estimated effect size = needs larger trial

-Did they reach their goal recruitment? If not, study may be underpowered (and may make an interesting discussion point on why not)

**Intervention (Facilitator);**

-Were the groups similar at the start of the trial, or did randomization “fail”? What do we do if the randomization fails (which can happen, just by chance). Often times control group can reflect out of date standard of care. Both groups may get more attention than std of care. Any possible difference in treatment aside from the intervention could explain any differences in the results

**Outcomes (Facilitators):**

-Why do trialists used composites? to increase the frequency that the outcome of interests occurs = smaller sample size required. However, have to be careful that the composites are either a.) similarly important, or relatively balanced in how often they occur – if many of the events were a less clinically meaningful results (e.g. NSTEMI as apposed to cardiac death), the result of the trial may not be as important

-Were the secondary analyses results consistent with what would be expected based on the primary outcome? Why or why not? (this may influence what mechanism we suspect to lead to the tiral result)

-When and why is blinding important? It’s ideal if the study is ‘double-blinded’ – that is, both patients and investigators are unaware of the treatment allocation. If the outcome is objective (death) then blinding is less critical. If the out­come is subjective then blinding of the outcome assessor is critical (so they don’t bias the data collection). Ideally, patients (placebo), doctors (diff in tx delivered), data collectors (bias in data collection), adjudictors of outcome (bias in outcome determination) should all be blinded

-How was the treatment exposure analyzed? Intention to treat analysis (= by assigned group is often best), because crossover will bias the study toward a smaller estimate of effect size if ITT is used = less false positive results

**Results (Facilitator):**

-How large was the treatment effect in absolute terms? How often did the primary outcome occur in the control group - this tells you about Low or high-risk cohort

-Were all randomized patient data analyzed? If not, was a sensitivity or “worst case scenario” analysis done?

-Was there much crossover between groups? What effect might this have on the trial result? In ITT analysis, it may make estimate of effect size understated. If not ITT, may bias the result in either direction = increases the risk of a false positive finding.

-Was there significant loss to follow-up? – if higher than 20% (arbitrary threshold commonly used in the lit) was a reason given why? Can bias the results (e.g. if a treatment is poorly tolerated, those patients are not included in the final analysis). There will always be some loss to follow-up, even in a well done trial. Why does loss to followup threaten validity? (if it is totally random who follows up, it doesn’t – but usually it is not random, and can bias the results if patients with particular outcomes are more or less likely to followup. Intention to treat can help ameliorate this and cross-over)

-What methods did they use for dealing with missing data? (carry forward last data point, vs single imputation, vs multiple imputation)

-Was the statistical power sufficient to draw conclusion about the subgroup analyses? When looking at a subgroup, there is a smaller n, and less power. So often interesting questions can’t be answered without another study.

-Are the results of the secondary analyses consistent with what would be expected based on the primary outcome result? Why or why not?

-Often, there are many adverse effects in the placebo group – why?

-Why might relative risks (as opposed to absolute) be important to report? They are less influenced by the baseline risk of the population – and so may represent the effect size of the treatment better – but are less applicable to patient care. The lower the risk in the control group, the larger the difference between relative risk reduction and absolute risk reduction.

-P = 0.05 – did the authors heavily inform their conclusion based on this threshold of significance? Were any of the non “statistically significant” findings potentially clinically important? Keep in mind that “absence of evidence” is not equivalent to “to evidence of absence”, and thus that findings that do not reach statistical significance for a difference should not be considered equal (this is a common error in manuscripts).

**Critique (Facilitator):**

Threats to internal validity:

-Were there conflicts of interest: What was the funding source, if they say? Intellectual conflicts of interest of the authors?

-Were there any large deviations from the pre-registration plan? Are these important – how so?

-Discussion point: what might cause the trialists to change design or statistics during the trial? Internal (slow recruitment, better estimate of effect size) vs external factors (e.g. loss of equipoise with contemporary trails)

-Did they deviate from the EQUATOR network guideline for what should be reported in a clinical trial = CONSORT (<http://www.equator-network.org/reporting-guidelines/consort/>)? Almost all major journals have agreed that reports of RCTs should include these details in RCT reports

-Was the trial stopped early? If so, why? (and how – independent data monitoring committees usually make the decision… futility, loss of equipoise / clear benefit)

Summary of threats to internal validity from Dartmouth:

A screenshot of a newspaper

Description automatically generated

Threats to external validity:

1. -“ Is my patient so different to those in the study that the results cannot apply?” Many discussed in patients included section: representative-ness of recruited population (sampling bias, or if the trial was designed to interrogate a different population than we are interested in.)
2. -Is this treatment feasible in my setting?

Lancet 2005; 365: 82–93

|  |  |
| --- | --- |
| A screenshot of a cell phone  Description automatically generated | A screenshot of a cell phone  Description automatically generated |

**Application Facilitators Guide:**

-Is the demonstrated benefit large enough to warrant any adverse effects / costs associated with the intervention (which is almost always going to depend largely on what the patient’s baseline risk is for the outcome we’re trying to prevent)?

-Is this tradeoff likely to be consistent across patient’s with varying beliefs and preferences? Were the outcomes measured important to patients?

1. -Do we believe the results are true? (Internal Validity: overall assessment of the point brought up above)
2. -Do we believe the results are applicable to this patient? (Does the patient have differences in attributes from the patients enrolled in the trial that you might expect to influence treatment efficacy?)

--Do you expect the efficacy (treatment effect under idealized study conditions) and effectiveness (treatment effect in the real world) to differ? – e.g. if aggressive followup, expertise driven interventions, costs/adherence are likely to be an issue?