Therapy
Sample CAT worksheet for a therapy paper, from McMaster University (go ahead and
download the paper version, which has answers and sample calculations):
Therapy Paper Worksheet

Key concepts:
General.
Therapy papers are not exclusively evaluations of single drugs; for example, treating
blood pressure in diabetics is “therapy”. Likewise, drugs may not be involved at all; for
example, banding of varices for acute GI bleeding in cirrhotics is therapy.

Are the results valid?
Æ Validity of a study is simply closeness to the truth; ask yourself: “Can I believe the
results?”
Æ Validity is not “all or none”; for any given study, your needle will end up pointing
somewhere between 0% and 100% valid. It’s not a yes or no question.
Æ The more bias that is introduced into a study, the less valid (less close to the truth) the
study becomes.
Æ Ways to limit bias:
   ♠ Randomize patients to treatment and control groups. This accounts for both
known and unknown prognostic factors at baseline.
   Pearl: If randomization occurs, then p values are not necessary in the
   traditional “table 1” baseline comparison of the two groups. p values
   measure the probability that differences between the groups occurred due
to chance (and if they were randomized, then we know any differences are
from chance, thus rendering p values unhelpful!)
   ♠ Perform an intention to treat analysis. That is, analyze patients based on the
   group they were initially randomized to, not the one they end up in. This
   preserves the initial randomization by keeping known and unknown prognostic
   factors equally distributed between the two groups throughout the study.
   ♥ Blind patients to which group they are in. Otherwise, those who know they are
   in the treatment arm may report better outcomes because of the placebo effect
   (more important for subjective outcomes, such as quality of life scores).
   ♦ Blind clinicians to the group allocation. Clinicians may consciously or
   unconsciously treat the two groups differently if they are not blinded.
   ♠ Blind those assessing outcomes to the group allocation. If the outcome is all
   cause mortality, this binding is less important; but if the outcome is at all
   subjective (what was the cause of death?), this blinding becomes very important.
   ♥ The term “double blind” is useless; you want to know exactly who were and
   were not blinded to the group allocation (Patients? Clinicians? Outcomes
   assessors?)
Æ Losing a great proportion of patients to follow up may compromise validity.
   Pearl: Do a worst case analysis (for example; assume all patients lost in the
treatment group died, and all those lost in the control group lived) and see if it
   significantly changes the stated result of the trial. If so, validity has been
   compromised.
**What are the results?**

→ Some experts argue that the most important outcome is always all-cause mortality. However, practically, we are often interested in other outcomes such as first ever heart attack (primary prevention), repeat heart attack (secondary prevention), stroke, progression of renal disease, quality of life, days in the hospital, and on and on.

→ Beware of composite endpoints! This is becoming epidemic in cardiology (and other) trials. A composite endpoint of death, MI, and repeat catheterization is not useful if a significant difference is driven entirely by decreased rate of repeat catheterization!

→ How large is the treatment effect? This is commonly measured by relative risk and its corollary, relative risk reduction. Why? Because these are bigger numbers than absolute risk and absolute risk reduction, and thus “look better”.

→ Keep in mind that if the treatment arm has a 0.01 (1%) absolute risk of developing kuru, and the control group has a 0.03 (3%) absolute risk of developing kuru, then the relative risk reduction is a whopping 66%! (Whereas the absolute risk reduction is 0.02, or 2 %) The bigger number almost always ends up being the one reported: Our drug decreases the risk of kuru by 66%! vs. Our drug decreases the risk of kuru by 2%!

 Pearl: If the type of risk reduction is not specified, always assume it is the relative risk reduction being reported.

→ Formulae:

<table>
<thead>
<tr>
<th></th>
<th>Died</th>
<th>Survived</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirin</td>
<td>40 (a)</td>
<td>560 (b)</td>
</tr>
<tr>
<td>Placebo</td>
<td>60 (c)</td>
<td>540 (d)</td>
</tr>
</tbody>
</table>

CER = Control Event Rate  
EER = Experimental Event Rate

- \( \frac{c}{c+d} = \frac{60}{60+540} = 0.1 = 10\% \)
- \( \frac{a}{a+b} = \frac{40}{40+560} = 0.066 = 6.66\% \)
- \( RR = \text{Relative Risk of dying} = \frac{EER}{CER} = \frac{0.066}{0.1} = 0.66 = 66\% \)
- \( RRR = \text{Relative Risk Reduction from ASA} = 1 - RR = 1 - 0.66 = 0.34 = 34\% \)
- \( ARR = \text{Absolute Risk Reduction} = CER - EER = 0.1 - 0.066 = 0.034 = 3.4\% \)
- \( NNT = \text{Number Needed to Treat} = \frac{1}{ARR} = \frac{1}{0.034} = 29 \)
- Relative Odds (Odds Ratio) = \( \frac{a/b}{c/d} = \frac{(ad)}{(cb)} = 0.64 = 64\% \) *

*(see below in the Harm section for further discussion of odds and risk)

→ How precise is the estimate of the treatment effect? The true risk reduction can never be known. The best we can do is come up with a *point estimate*. For example, consider what would happen if you conducted a trial of ASA vs. placebo, and came up with an ARR for death of 3.4%. Then, let’s say a month later you repeated the study with a different group of patients. Do you think the ARR for death would once again be 3.4%? What if you repeated the study 100 times, would you get 3.4% each time? Unlikely!

→ We then place that point estimate within a neighborhood in which the true value is likely to fall: the *confidence intervals*. The larger the study, the narrower the confidence intervals will be. In the above example, the 95% confidence interval might be 1.2%-6.9%, with the point estimate being 3.4%.

→ We feel comfortable that ASA is actually preventing death when our 95% confidence intervals do not cross 0.0%. If that happens, then we can’t be confident that ASA is actually preventing death at all!
Using the above example, you can state: “I am 95% sure that the true treatment benefit of ASA lies between 1.2% and 6.9% ARR”.

The more confident you want to be, the broader the C.I. range will be. For example, if you wanted to be 98% certain that the true value lies between your confidence intervals, the range might be 0.07% to 20.3%.

What is a p value? Simply, the probability that the observed difference is due to chance. If the p value for ASA vs. placebo is 0.04, then we can make the statement: “I am 96% sure that ASA is preventing death compared to placebo. But, I acknowledge that there is a 4% possibility that the difference I see is due to chance, and that ASA is no better than placebo”.

Pearl: Assigning statistical significance to p values of <0.05 and C.I.’s of 95% is purely arbitrary, but widely accepted. Ask yourself, is a p value of 0.06 not helpful?