Pilot Studies: The Limits of Reality

Career Development Seminar

John E. Connett, PhD
Director, BDAC
June 24, 2014

Pilot Studies: The Limits of Reality

What's Wrong with Pilot Studies?

- Answer 1: Nothing …
- Answer 2: They can't do everything
- Answer 3: First ask: what SHOULD they do?
- Answer 4: Then ask: what CAN they do?
- Answer 5: Then go back and ask: what do I WANT them to do?

Pilot Studies: Why Do Them?

- Because you have a Big Dial Question, but you need some data to put in the application (for an R01, P50, etc)
- You need to develop a protocol
- You need to estimate outcome parameters (percent success, change in serum creatinine, change in QOL, etc)
- Your statistician says: what's the variability of your measurements?

Digression on Variability …

- Frequently you know the variability of one-time cross-sectional measurements: e.g., variability between people of diastolic blood pressure.
- But your primary outcome is CHANGE in the measured outcome across, say, a 6-month time period. What you really need, and what you often do not have – without a pilot study - is within-person variability

What You CAN'T Do in a Pilot Study

- Answer a Big Dial question that would require a Phase III clinical trial – because:
  - You don't have access to enough patients
  - You won't have enough funding
  - You don't know what dose to use or what protocol might actually work
  - Your statistician says: you need more data before you can write a proposal for the Big Study

Big Dial vs. Little Dial

Big Dial: Important, includes red zone for danger

Little Dial: Not crucial, no critical zone

If you are in a submarine, which dial do you watch?

Mistakes That Pilot Proposers Make

- Overestimate recruitment
- Propose to test a treatment with an implausibly large treatment effect (e.g.: improve survival in acute myelocytic leukemia from 30% to 60% …)
- Forget about the intention-to-treat principle
- Rely on ‘historical data’
- Have no estimate of variability of the proposed outcome

Reviewers Say:

- Can you demonstrate that you can recruit enough patients?
- What's your justification for your treatment effect? Is it plausible?
- What's the variability of your outcome?
- Can you retain patients?
- What is your rate of missing data, and what are you going to do about it?

You Are Shot Down

What SHOULD You Be Doing in a Pilot Study?

FEASIBILITY

- A Pilot Study is intended to show that you are capable of doing a full-scale study, and to provide data that you can use to justify a proposal for the Big Dial study that you really want to do.

Elements of Feasibility

- Can you get enter enough eligible patients? How many refuse consent? Do you have too many eligibility requirements?
- How many patients do you have to screen to get 1 patient in the study? If you have screen 1,000 patients to get 20 eligible and consented, how generalizable are your findings?
- Can you RETAIN patients? [Hint: a 15% rate of loss to follow-up is a RED FLAG]

More Elements of Feasibility

- Can you carry out your proposed protocol?

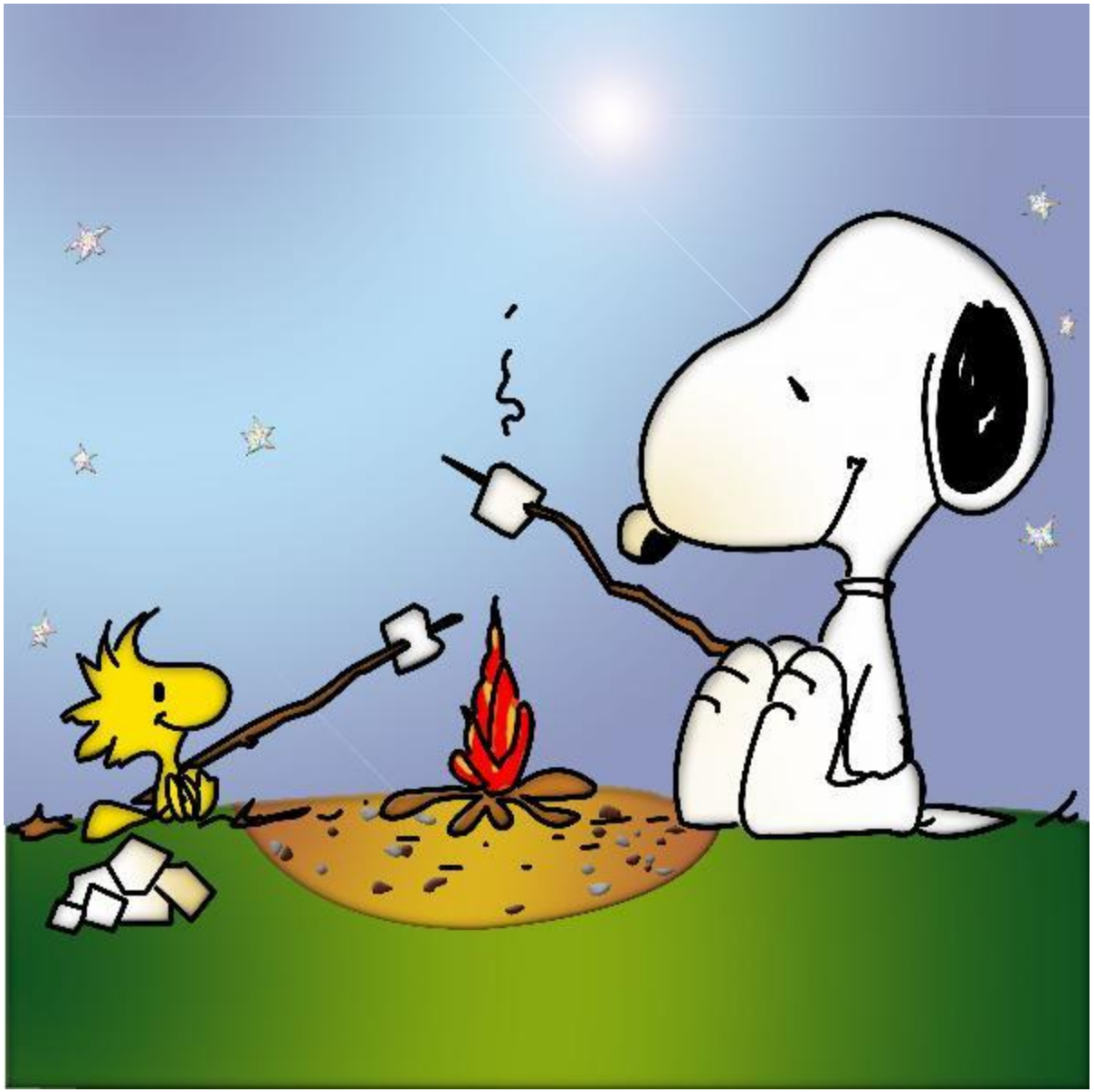
  E.g., can you collect induced sputum to evaluate microbiota in the lung?
- Can you estimate costs?
- Can you estimate what the outcomes are for each treatment group?
- Can you estimate rates of serious adverse events?

More Elements of Feasibility

- What is an optimal dose? (Efficacy vs. side effects …) Will your Pilot Study enable you to estimate it?
- Can you develop a good Manual of Operations?
- After completing the Pilot Study, will you have a workable protocol that could be implemented in a full-scale study?

Elements of Feasibility, Contin.

- Can you demonstrate that you can put together the resources and the staff to carry out a full-scale study?



Digression on Intention-to-Treat

- The Intention-to-treat principle comes up especially in Phase III clinical trials. Basically it says, everybody who is randomized is entered into the analysis as being in the group to which they were randomized, EVEN IF:

  - They move to Estonia and are never seen again

  - They never comply with their assigned treatment and perhaps even switch to the opposite treatment

* High-end journals (*NEJM, JAMA*, others) will INSIST on intention-to-treat analyses for Phase III clinical trials. Not so much for Pilot Studies, but you need to bear it in mind when you use Pilot data to plan a larger study.

Dropouts and Missing Data

- In any study, pilot or full-scale, you MUST have a plan for avoiding dropout.
  - Stopping use of the assigned medication should NOT be counted as dropout, as long as you can still evaluate the primary outcome
  - Dropout implies missing data. Missing data is almost NEVER missing at random. Which means, if dropout occurs, the results are probably BIASED. Best advice for handling missing data: prevent it!!!
  - BIAS is a four-letter word.

Retention, etc.

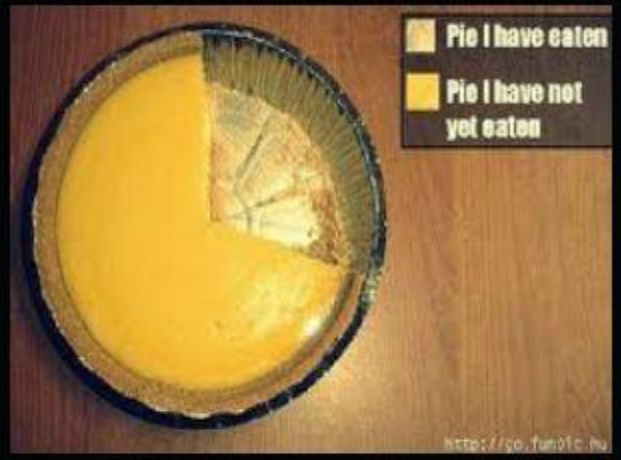
- Here is a good way to make a statistical reviewer SEE RED. Say that you expect a 15% rate of noncompliance with your experimental drug, therefore you will REPLACE noncompliers with additional people to achieve the original sample size.

  Which means, you keep replacing people until all you have left to evaluate are the good compliers.

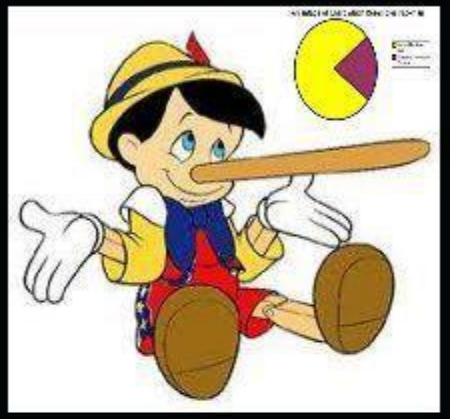
  Result: BIAS. Four-letter word! NOT consistent with intention-to-treat

STATISTICIAN

What my friends think I do



What my mom thinks I do



What society thinks I do

What my boss thinks I do

What I think I do

What I actually do

How to Work With a Statistician

- Statisticians are sensitive, caring souls who actually want to help you! Though this may not be obvious at first …
- DO contact the statistician well in advance of the due date! Two weeks is cutting it very close!
- DO involve the statistician at the very beginning, before you have a study design, possibly even before you know what question you want to ask
- You need to arrive at an unambiguous, clinically meaningful outcome for your study

How to Work with a Statistician contin.

- Make it clear whether this is a pilot study or a later-phase study. If the latter, you need some idea of:
  - What the treatment effects might be
  - What the variability of the outcome is
  - Or, if the outcome is dichotomous, what you estimate the percents of success or failure to be
  - How many groups (or doses, etc.) you plan to study, and why

Objections to a Focus on Feasibility

- But Sir, reviewers expect us have specific hypotheses, and we have to have statistical power to address those hypotheses …
  - If you had sufficient statistical power, you wouldn't be proposing a pilot. You would be proposing a full-scale study
  - You can always put on the appearance of having sufficient statistical power by proposing an alternative that is totally implausible. Is that a good idea?
  - You can propose a modest alternative that is plausible, but will require 100 times as many patients as you have access to

More Objections to a Focus on Feasibility

- But Sir, how can I explain this to reviewers?
  - First, state what your Big Dial Study would be: your vision of a large, definitive study which could prove that your New Idea for preventing Alzheimer's has merit
  - Second, state that you have to have some solid, reliable data on recruitment, estimates of treatment effects and variability of outcomes, retention rates, proof that the protocol can be carried out
  - Third, state that carrying out a pilot is the only sure way to get sound estimates of costs for a full-scale study
  - Fourth, describe how results of your Pilot would be used to inform the design of the Big Dial Study

What About Sample Size?

- But Sir, reviewers want to know what our sample size will be for the pilot study. If we tell them our sample size, then they want us to justify it!

  - True enough. Typically sample size for pilot studies is small. In many cases, it is an estimate of how many patients you have access to in a 1-3 year study. You have to argue that this number will be sufficient to address questions regarding recruitment, retention, treatment effect and variability.

Sample Size, Contin.

- For example: you do a pilot of Drug A. You know that the standard treatment, Drug S, is 50% effective. You want to estimate the effect of Drug A. Say you think it might be effective in 60% of cases. Your goal is to find 90% confidence limits for the effect. If you would like the 90% confidence interval to be about (45%, 75%), you will need about 29 people in the Drug A group. This interval is very wide, but it might give you a basis for planning a larger, definitive study.

More Objections Regarding a Focus on Feasibility

- But Sire, reviewers are interested in efficacy for pilot studies; they are less interested in effectiveness …

  Definition, efficacy: Does the treatment work in people that religiously take their pills?

  Definition, effectiveness: Does PRESCRIBING the treatment work, allowing for noncompliance, dropout, etc.?

There is some truth to this. However, planning studies strictly on efficacy frequently results in unrealistic, nongeneralizable studies.

Some Compromise Approaches

- You realize you don't have a large enough sample size to have adequate power to address your main question. But you want to say SOMETHING about the probability of seeing a trend. Here are a couple of options:
- Set the significance level at something bigger than the usual 0.05. Say, e.g., alpha = 0.15. What this means is, if at the end of your study, you find a difference significant at the 0.15 level, you will consider that strong enough evidence to proceed with planning a larger definitive study

Compromise Approaches, Contin.

- Instead of computing the probability of obtaining a significant result (given a specific alternative hypothesis), compute the probability of seeing a positive trend. Your power for detecting the alternative hypothesis may be 0.50 (not usually considered adequate power), but your probability of seeing a positive trend if there really is one may be 0.95.

Caveat Emptor …

- The compromise approaches just mentioned are not universally accepted. So if you are going to propose these for you pilot study, you will need to be VERY CLEAR about it. Remember, your statistician-reviewer may be slightly less caring and sensitive than your BDAC collaborator!

The Happy Pilot

Questions, Comments?