Using Statistical Thinking to Optimize Scientific Studies

Kristi L. Griffiths, PhD
Rebecca Elliott
Laura L. Scheppers

Eli Lilly & Co.

Quality & Productivity Research Conference
June 4-6, 2007 Santa Fe, NM

Answers That Matter.

Outline

- Overview
  - Abstract highlights
  - Example
  - Common issues with scientific studies
- Elements of statistical thinking in study design
  - Articulating the question
    - Illustration
  - Designing the study
    - Proper replication
- Practical solution

  PLAN → DO → ANALYZE

Abstract highlights

Across all industries today, the need for quick answers may result in **haphazard, inefficient, and less effective** experimental studies. There are several issues that poor studies have in common, e.g.

- ✓ **a vaguely articulated research question,**
- ✓ **no specific idea of what information to gather,**
- ✓ **insufficient/inefficient study design,**
- ✓ **improper execution and/or data collection,**
- ✓ **inability to analyze/interpret the data, and/or**
- ✓ **an inability to turn raw data into useful information**.

Example

“Do you have a couple of minutes for a quick question? I did two runs of Process A vs. Process B and measured the yield. Looking at the numbers, I’m positive that Process B gives much better yield. What is the best way to express this with some statistics?”

Example

"Do you have a couple of minutes for a **quick question**? I did **two runs** of Process A vs. Process B and measured the yield. Looking at the numbers, **I'm positive** that Process B gives much better yield. What is the best way to express this with some **statistics**?"

Immediate thoughts:

Was this study designed in the first place?

Is this the end of the study?

→ What question are we trying to answer?

Common Issues with Scientific Studies

Replication? N=2 process runs?
What does N=2 mean?

Independence? How?

Confounding?
Anything not mentioned that I should know about?

And … *what was the question to be answered?*

Elements of Statistical Thinking in Study Design

The approach to good study design can be illustrated as:

1. Articulate the question
2. Thoughtfully & deliberately design the study
3. Gather data according to the study design
4. Analyze the data according to the study design
5. Make a decision

What does this resemble?

The Scientific Method

Elements of Statistical Thinking in Study Design

Scientific Method

Hypothesis:	Ask a question, form a hypothesis
Design:	Design an experiment to answer the question
Data:	Conduct the experiment, gather data
Conclusion:	Analyze results, draw a conclusion, see if it supports or refutes the hypothesis

(Repeat)

Elements of Statistical Thinking in Study Design

Scientific Method

Hypothesis: **Ask a question**, form a hypothesis

Design: **Design an experiment** to answer the question

Data: Conduct the experiment, gather data

Conclusion: Analyze results, draw a conclusion, see if it supports or refutes the hypothesis

(Repeat)

Articulating the question

The question drives the experimentation!

A.K.A.

If you don’t know where you’re headed,
any path will take you there!

It is critical to have a clearly defined question!

Articulating the Question: Example

Q1: Is Process B better than Process A?

With respect to yield? Cost? Processing time? …..what?

Q2: Does Process B make widgets more quickly than Process A?

How "quick" does Process B need to be? What's the measure of success?

Q3: Is Process B at least 2 hours faster than Process A?

OK, we're getting there

Elements of a Refined Question

1. Includes a specific, measurable attribute

   e.g. average, standard deviation, every individual result, etc.

2. States the criterion for success

   e.g. difference must exceed 2 hours,

   all results must be within 3% of target,

   the variability must be cut in half

3. Refers to the inference space

   e.g. for all lots manufactured by process B,

   all lots made by process B at site Z during the rainy season

   -- helps to define the required replication strategy

Refining the Question

Recall:

Q3: Is Process B at least 2 hours faster than Process A?

OK

Q4: Is Process B expected to be at least 2 hours faster on average than Process A for all future lots.

Refined

Illustration: Defective Tablets

Poor Does the type of raw material A have an effect on tablet quality?
Good Does the **percent of defective tablets** within a product lot differ across two types of raw material A?
Refined **For all future tablet lots**, does the **percent of defective tablets** **differ by more than 3%** depending on the type of raw material A?

Designing the Study

1. Articulate the question
2. Thoughtfully & deliberately design the study
   (to answer the question)

Need to design the study so that:

- The study structure takes into account
  - **Proper Replication**
  - Confounding
  - Randomization
- The “rules” for collecting samples and generating / analyzing / interpreting data are clear

Designing the Study: Proper Replication

Elements of a Refined Question:

3. Refers to the inference space

   -- helps to define the required replication strategy

Example: The dilution scheme in a chemical assay was suspected to contribute a significant amount of variability to the method. The scientist wants to know if the variability of the assay increases by more than 20% when changing from dilution scheme A to dilution scheme B. The scientist conducted 15 dilutions of each scheme and assayed the material on one run.

Did the scientist conduct "proper replication" to answer the question?

We are interested in learning if there is a difference between dilution schemes.
YES, the multiple executions of each dilution scheme is proper replication.

Practical Solution

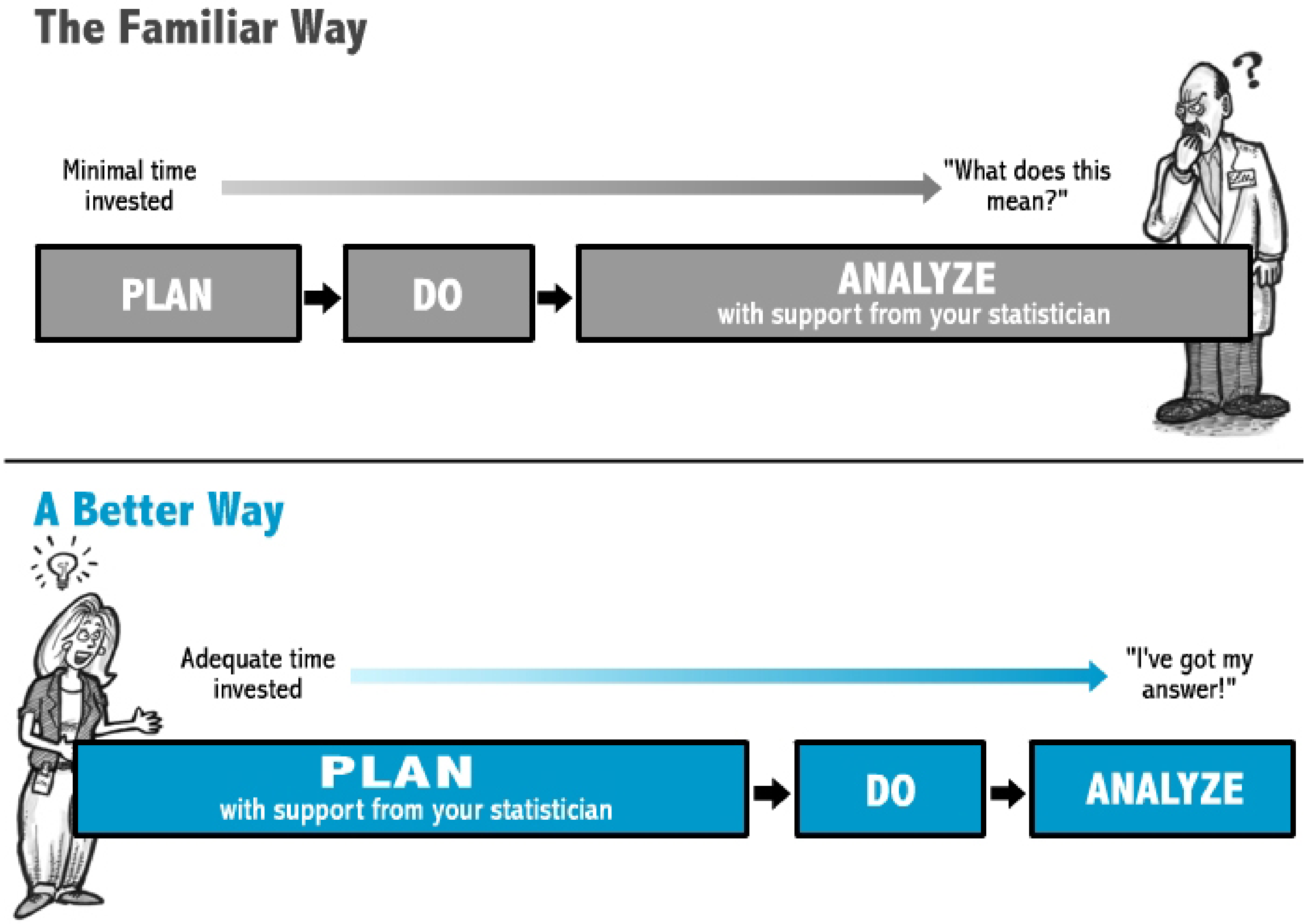
PLAN → DO → ANALYZE

But, where in this process should the effort mainly occur?

HINT: ***"If you think there isn't time to do it right, when do you think there will be time to do it over?"***

- ✓ Need to balance speed of the effort with the quality of the science
- ✓ Activity gives the ***appearance*** of progress
- ✓ Better to make thoughtful and deliberate decisions than to have to augment / re-do / scrap the work

PLAN → DO → ANALYZE



Summary

Given the urgency of development / innovation, it is difficult to take a methodical and thoughtful approach to experimentation.

Balancing the required speed with the desired quality of decision making, it is imperative (and more efficient) to follow the scientific method.

- **Articulate the question**

The question drives the experimentation !!!

Acknowledgments

Laura Scheppers
Pharmaceutical Delivery Systems Statistics

Gwendolyn Cook
HR – Training/Development