White Paper

Use Statistical Thinking and Methods to Make Better Decisions

Dr. Ronald D. Snee, Ph.D., Founder and President, Snee Associates, LLC

The scientific method is fundamental to the operation and success of the pharmaceutical industry including the discovery of new products, development of the products using clinical trials and the manufacture of the products. The scientific method has three key components:

- Recognition and formulation of a problem
- Collection of data through observation and experiment
- Formulation and testing of hypotheses

Clearly collection of data is central to the scientific method. This is where statistical analysis comes into play. First is the design and implementation of the data collection process. Then there is the analysis and interpretation of the data.

So why are data important? Why do we need data?
Lord Kelvin, the creator of the Kelvin temperature scale admonishes:

"When we can measure what we are speaking about and express it in numbers, we know something about it, but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind."

Data provide knowledge. Data provide an historical record of the events of all types. Human memory fades and information gleaned from observation is lost. The data provides information on what happened. In my experience, discussions without data typically produce more heat than light.

Data quality is also important. Data provide the information that are needed to solve the problem being studied. Data collection using statistically designed experiments typically produces high quality data. Sometimes we are faced with observational data: data collected without the aid of a statistical design. Examples include batch records, raw material characteristics and product characteristics data such as tablet weight, thickness, hardness and dissolution. Table 1 shows some limitations of observational data.

Table 1: Limitations Typically found with Observational Data.

Table 1
• Data records are frequently incomplete. • Predictor variables may have missing values. • Important predictor variables may be omitted. • Data are not collected at proper frequency, e.g. Weekly when daily data are needed. • Predictor variable varied over a small range reduces the probably of finding an effect. • Correlated predictor variables that hinder the identification of important variables.

Now that we have the data, or even before, we see people asking what statistical tools we should use? A better question to ask could be what is the problem that we are trying to solve. Once we know the problem, the statistical tools to use will become readily apparent.

Statistical Significance versus Practical Significance
When using statistical analysis, there is always the question of p-values, statistical significance, and how it compares to practical significance. They are not the same. Statistical significance tells us whether the observed effect or difference is larger than can be attributed to random chance variation. An effect or difference may be statistically significant but be so small as to be of no practical significance. In fact, with a very large sample size, a small effect may be statistically significant and have no practical significance.

The best way to think about practical significance is to determine whether the effect/difference results in a change in behavior. Will we switch suppliers? Do we need to set the process temperature at a different level? Do we need to scrap the product? If we don't do something different then the effect is of no practical importance.

Importance of Graphics
Renowned Princeton professor John W. Tukey reminds that "The greatest value of a picture is when it forces us to notice what we never expected to see." While I knew the importance of graphics, an experience very early in my career gave me a lesson I will never forget.

I was working with a group that was developing a new product. Unfortunately, the corporate toxicology laboratory claimed that the product was toxic. I was asked to analyze the data and present my results at a meeting with the Director of the toxicology laboratory. I made my pitch, complete with all appropriate statistical analysis which showed that there was no evidence of product toxicity. As I was making my presentation, it was clear from the Director's body language that I was not making my point.

Just prior to this meeting, I had hand-drawn a small graph of the data. The graph was a dot plot of the results with the control and test groups shown side-by-side on the graph. When one looked at the difference between the groups in comparison to the variation within the groups, it was clear, as indicated by the statistical analysis, that there was no difference between the test and the control groups.

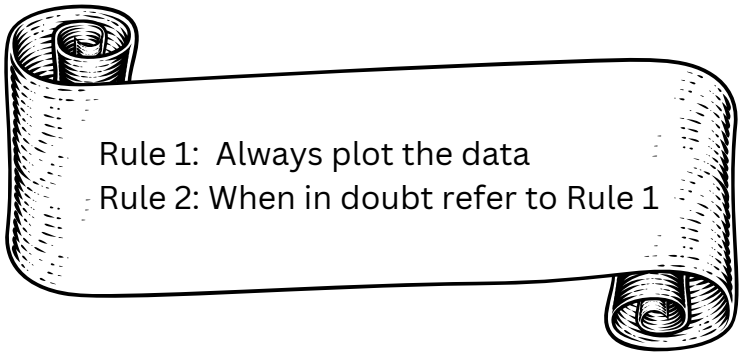
I had not shown this plot in my presentation, but I had it available. I passed the plot down the table to the Director. An interesting few minutes elapsed. The Director would look at the graph then make some comments regarding the study, then look at the graph again. This activity went on for about 5 to 10 minutes. Then finally he said to the group, "well I don't think we've demonstrated that there is a significant difference between the test and the control groups. I think that we need to do some more toxicity testing of this product".

This was a big win for me and my client. This further emphasized the importance of creating graphical display of data. Of course, in hindsight I should have started off the presentation with this graph, but for whatever reason, I didn't. I had learned an important lesson: a picture can be worth a thousand words.

Graphics Principles

Early in my career, I decided that for any written report or any presentation, it would be standard operating procedure for me to have some graphs to display the important findings. The graphics would encourage the audience or the reader to look immediately at the graphics. This would hopefully enable them to understand/agree with my findings and interpretation. I encourage students and all those I work with to adopt and utilize this operating principle. Figure 1 shows some rules of graphics in use. My professor at Rutgers university admonished his students to “Plot the data and think”!

Figure 1: The Golden Rules for Statistics



I conclude that using statistical thinking and methods can help us make better decisions. To be successful we need to think broadly considering the problem to be solved, the scientific method, data, graphical visualization of data and statistical tools.

Reference

Hoerl, R. W and R. D. Snee (2020) *Statistical Thinking - Improving Business Performance, 3rd Edition,* John Wiley and Sons, Hoboken, NJ.

Ronald D. Snee Bio

Ronald D. Snee, PhD is Founder of Snee Associates, LLC, a firm dedicated to the successful implementation of process and organizational improvement initiatives. He worked at DuPont for 24 years prior to starting his consulting career. He has served as Adjunct Professor in the pharmaceutical programs at Temple and Rutgers Universities.

Ron is an Honorary Member (Hall of Fame) of ASQ and has been awarded ASQ’s Shewhart, Grant and Distinguished Service Medals, and ASA’s Deming Lecture, Dixon Consulting and Hahn Quality and Productivity Achievement Awards. He is a frequent speaker and has published 10 books and more than 350 papers in the fields of statistics, quality, pharmaceuticals, performance improvement and management.