



# APPLICATION OF STATISTICAL TOOLS IN THE DEVELOPMENT OF PHARMACEUTICAL PRODUCTS

VIPUL PATEL <sup>(1)</sup>, DHAVAL PATEL <sup>(2)</sup>

<sup>(1)</sup> Research Scholar, Gujarat Technological University.

<sup>(2)</sup> Department of Pharmaceutics, Saraswati Institute of Pharmaceutical Science,  
Dhanap, Gandhinagar, Gujarat, India. 382321

## ABSTRACT

Industrial production procedures for pharmaceutical product development require a high level of expertise and control to prove that the end product will be of the requisite quality. Usually, a lot of information is acquired throughout the process however, in order to make use of the data, the information that is contained must be extracted using the appropriate statistical tools. Large quantities, varied sample rates, process noise, and non-linear connections are some of the difficulties encountered by industrial process data. To understand the effects of multidimensional and interactions of input parameters on the output responses of pharmaceutical goods and analytical methodologies, design of experiments (DoE) have been employed extensively. SPSS, STATA, Minitab, Graph pad Prism are some of the popular programs for statistical analysis. EXCEL and other spreadsheet programs are user-friendly and excellent for entering, coding and storing survey data. Larger-scale manufacturing problems might occasionally occur, and statistical optimization could provide solutions. Furthermore, it is applied in medical research to increase efficiency and precision. This article includes theoretical and practical approaches for using Design of Experiments (DoE), Minitab, IBM SPSS, Graphpad Prism in the pharmaceutical industry.

**KEYWORDS:** Statistical Tools, Minitab, IBM SPSS, Graphpad Prism, Design of Experiments (DOE)

## INTRODUCTION (1)(2)(3)

The evolution of the statistics and information management professions in the pharmaceutical industry has had two primary drivers. The first was the change from an industry that was fragmented and focused on the manufacture of nostrums and cure-ails compared with the cutting edge, highly scientific, research-oriented industry that exists today. The second was the growth of governmental control over the pharmaceutical industry in the latter half of the twentieth century.

There was a time when validating experiments through data was done fully using manual computation. This opened up gaps for human error and greater cost of carrying out research especially when the data was large by say, over 1000 field observations. Today, due to the gradual advancement in technology, statistical tools are being used in medical research for greater efficiency and accuracy.

Industrial statistics has played a key role in the creation of competitiveness in a wide range of organizations in the industrial sector, in services, health care, government and educational systems. The tools and concepts

of industrial statistics have to be viewed in the context of their applications. These applications are greatly affected by management style and organizational culture.

Industrial organizations typically include units dedicated to product development, manufacturing, marketing, finance, human resources, purchasing, sales, quality assurance and after-sales support. Industrial statistics is used to resolve problems in each one of these functional units. Marketing personnel determine customer requirements and measure levels of customer satisfaction using surveys and focus groups. Sales are responsible for providing forecasts to purchasing and manufacturing. Purchasing specialists analyze world trends in quality and prices of raw materials so that they can optimize costs and delivery time. Budgets are prepared by the finance department using forecasts that are validated periodically. Accounting experts rely on auditing and sampling methods to ascertain inventory levels and integrity of databases. Human resources personnel track data on absenteeism, turnover, overtime and training needs. They also conduct employee surveys and deploy performance appraisal systems. The quality departments commonly perform audits and quality tests, to determine and ensure the quality and reliability of products and services. Research and development engineers perform experiments to solve problems and improve products and processes. Finally, manufacturing personnel and process engineers design process controls for production operations using control charts and automation.

## 1. Statistical Tools Used in Pharmaceutical Industries

Researchers in the pharmaceutical sector and science frequently employ statistical methods. Many statistical tools and software are available for the research purpose and it can also be employed on industrial scale. Some of the statistical tools are listed below:

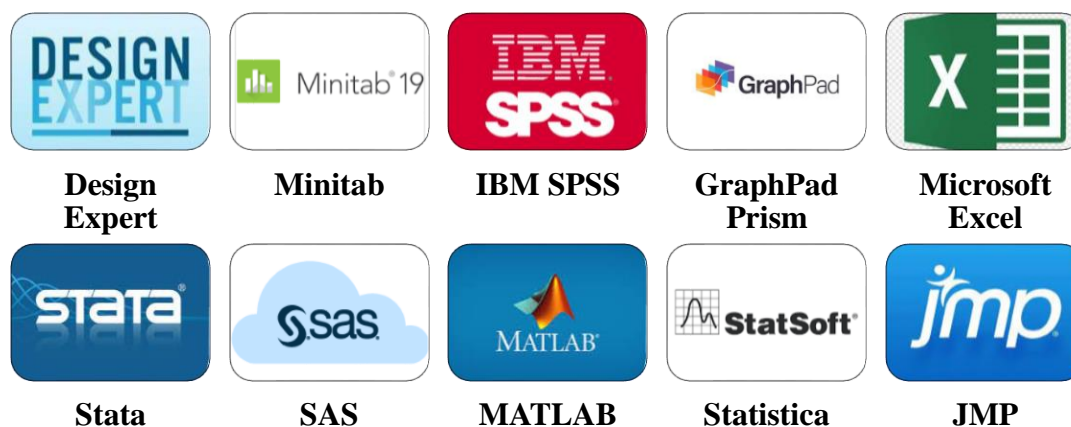


figure: 1 statistical tools used in pharmaceutical industries

### 2.1 Design Expert

#### 2.1.1 History (4)

Design Expert is a software from Stat Ease, which was begins in the 1970. Pat Whitcomb, Tryg Helseth and Mark Anderson mastered design of experiments (DOE) while working for General Mills Chemical Division's Process Development Group. When IBM came out with its first PC in 1981, Pat recognized an opportunity to make DOE easy for non-statisticians via a menu-driven computer program, which he named "Design-Ease" (DE).

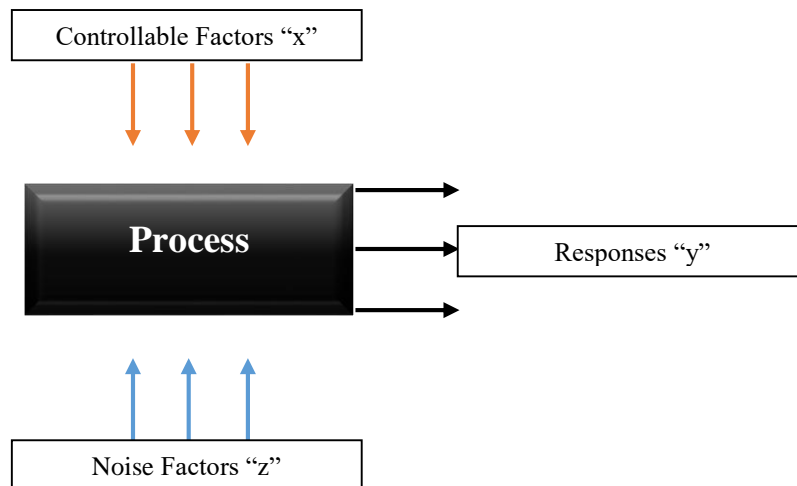
Pat then established "Stat-Ease" with the intriguing slogan "Statistics Made Easy" (an oxymoron?). In 1988, the company released its first version of Design-Expert (DX), which provided response surface methods (RSM) for process optimization. Moreover, DX offered mixture designs for formula optimization, which was very appealing to users in the process industries. In 1996, Stat-Ease incorporated all the features of Design-Ease into Design-Expert version 5 and translated it to Windows. Since it was so simple to use, Stat-Ease software allowed its experts, Pat, Mark, and one of their early hires, Shari Kraber, to offer computer-intensive workshops on DOE/RSM to thousands of academics over the years.

#### 2.1.2 Introduction (5)(6)

Design of experiment (DOE) is a planned approach for determining cause and effect relationships. It benefits from multifactor testing rather than evaluating cause and effect relationships. By applying DOE, we can also identify factor interactions.

Design Expert provides a full array of designs and analysis tools to address three overarching goals:

1. **Screen for vital few process factors or mixture components**
2. **Characterize critical interactions**
3. **Optimize for peak performance**



**figure: 2 design-of-experiments (doe) optimization of process input parameters to improve outcomes.**

Design Expert makes R&D easy with an intuitive interface and amazing graphics. It provides powerful tools to lay out an ideal experiment on process, mixture, or combination of factors and components. Build robust designs via in-line power calculations and the ability to add blocks and center points. Design-Expert makes it easy to see what, if anything, emerges as statistically significant and how to model the results most precisely. Automated model-reduction tools, paired with in-line diagnostic graphs, provide a streamlined analysis process. It provides the confidence need to present and publish findings. Design-Expert offers a wide selection of graphs that helps to identify standout effects and to visualize the results.

### 2.1.3 Why to use Design Expert? (7)

1. **Two-level factorial screening designs:** Identify the vital factors that affect your process or product so you can make breakthrough improvements.
2. **General factorial studies:** Discover the best combination of categorical factors, such as source versus type of raw material supply.
3. **Response surface methods (RSM):** Find the optimal process settings to achieve peak performance.
4. **Mixture design techniques:** Discover the ideal recipe for your product formulation.
5. **Combinations of process factors, mixture components, and categorical factors:** Mix your cake (with different ingredients) and bake it too!
6. Design-Expert program offers rotatable 3D plots to easily view response surfaces from all angles. Numerical optimization function finds maximum desirability for dozens of responses simultaneously!



figure: 3 start menu of design expert version 13

### Types of Designs in Design Expert

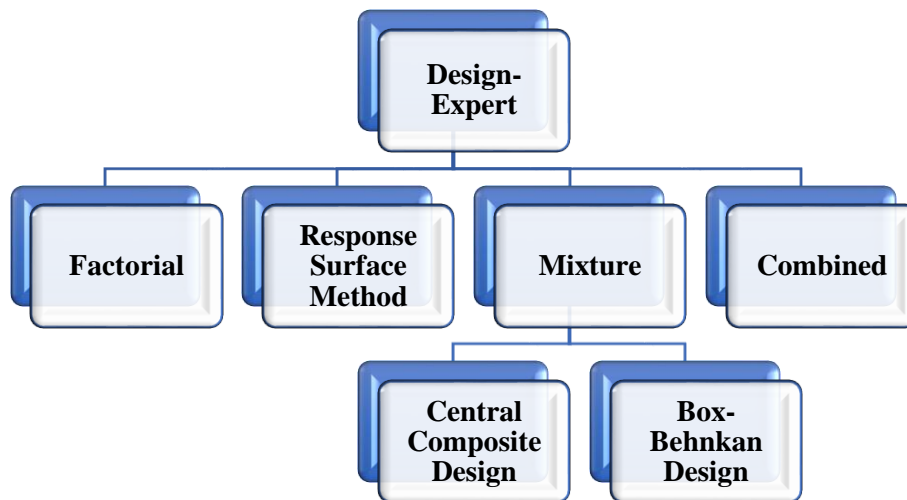


figure: 4 types of designs can be performed in design expert

## 1.1 Minitab

Working with large data sets, software can save considerable time and, in turn, the time savings and increased speed of analysis permits the analyst to pursue avenues of investigation that might not otherwise be possible.

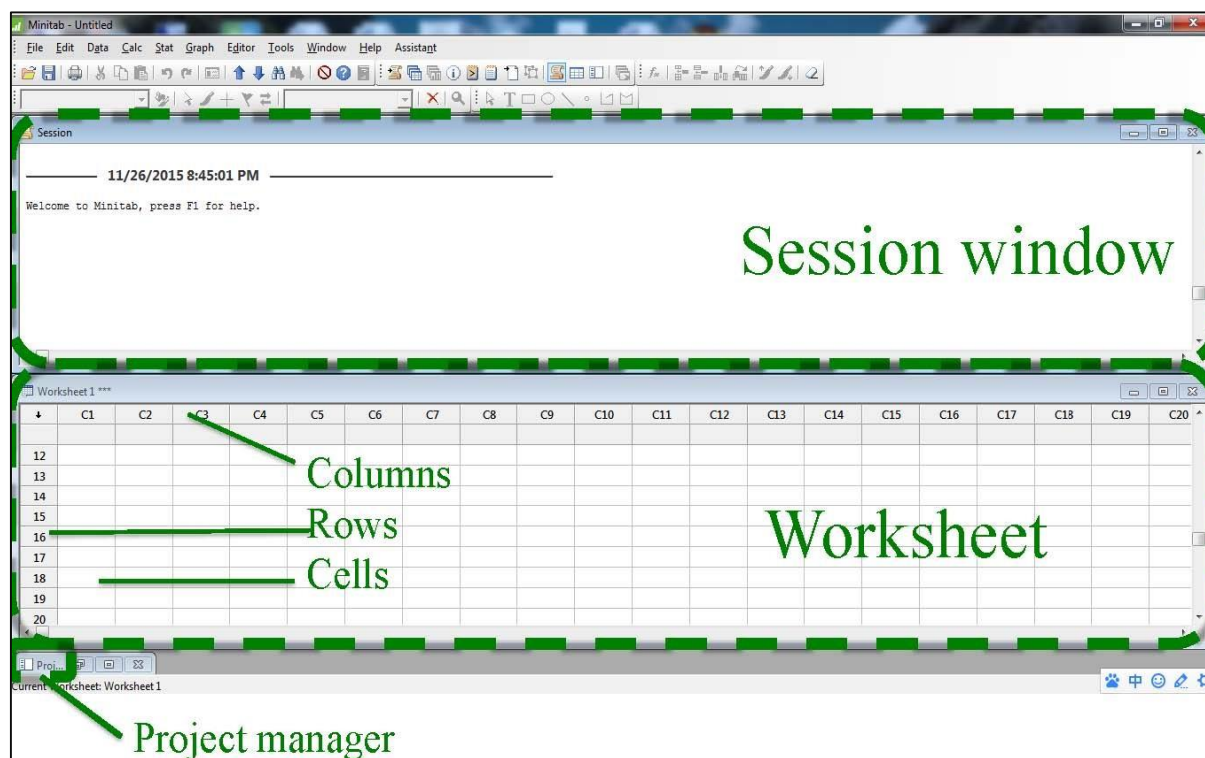


figure: 5 initial screen of minitab

MINITAB has two modes for submitting commands: a command line mode and a mouse-activated pull-down menu environment. Many people will find the mouse/menu environment easier to use. Most experienced MINITAB users are adept at both methods.

A good way to start learning how to use statistical software is to give an overview: see what it looks like, consider the different possibilities offered, enter some data, develop an idea on how to handle it and start doing some simple analysis.

## 2.3 IBM SPSS (8)

At Stanford University in the late 1960s, Norman H. Nie, C. Hadlai (Tex) Hull, and Dale H. Bent developed the original software system named Statistical Package for the Social Sciences (SPSS). They needed to analyze a large volume of social science data, so they wrote software to do it.

A statistic is calculated using a sample. In a sense, a sample is the keyhole you have to peer through to the population, which is what you're trying to understand. The job of SPSS is to calculate. Your job is to provide a good sample.

SPSS is comprehensive and flexible statistical analysis and data management solution. SPSS is a computer program used for survey authoring and deployment, data mining, text analytics, statistical analysis, and collaboration and deployment. SPSS can take data from almost any type of file and use them to generate tabulated reports, charts, and plots of distributions and trends, descriptive statistics, and conduct complex statistical analysis.

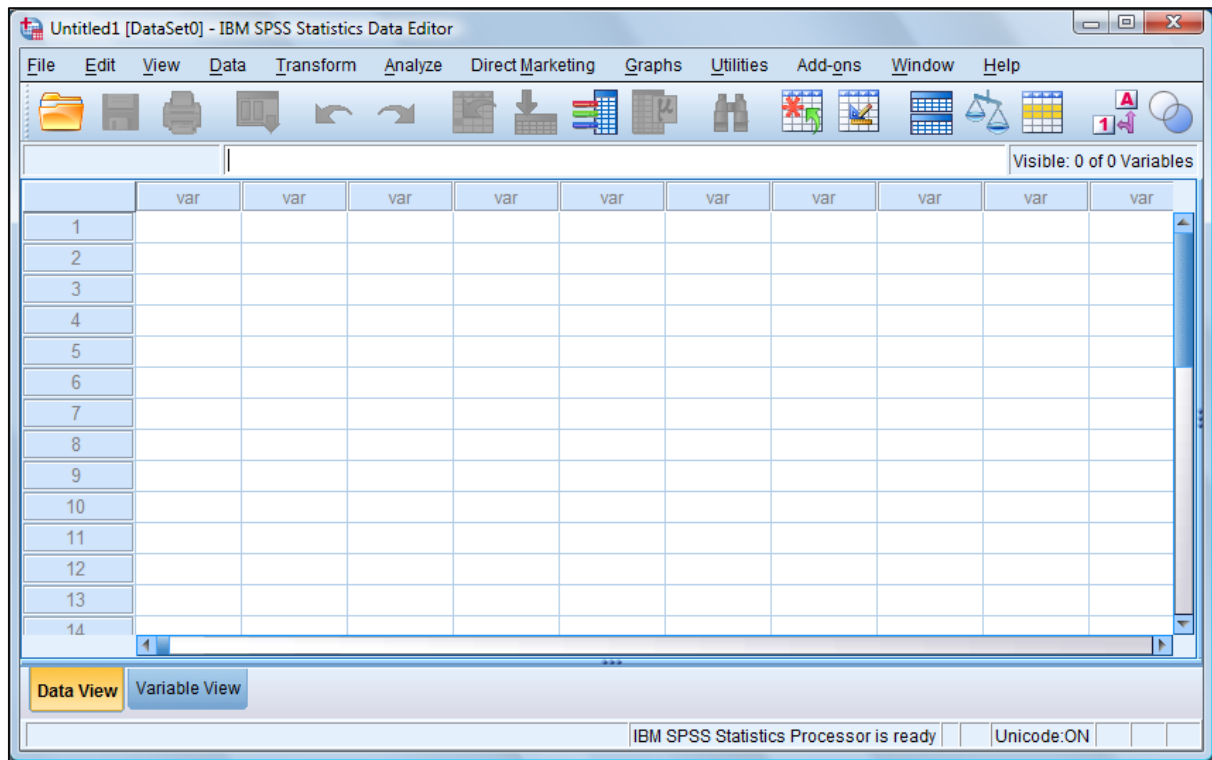


figure:6 initial screen of spss

### 2.3.1 Strengths of SPSS

1. Very robust statistical software.
2. Many complex statistical tests available.
3. Good “stats coach” helps with interpreting results.
4. Easily and quickly displays data table.
5. Can be explained.
  - a) Using the syntax feature.
  - b) Purchasing add-ins.

### 2.3.2 Limitations of SPSS

1. Can be expensive.
2. Not intuitive to use.
3. Typically requires additional training to maximize features (at a cost)
4. Graphing feature not simple as excel.

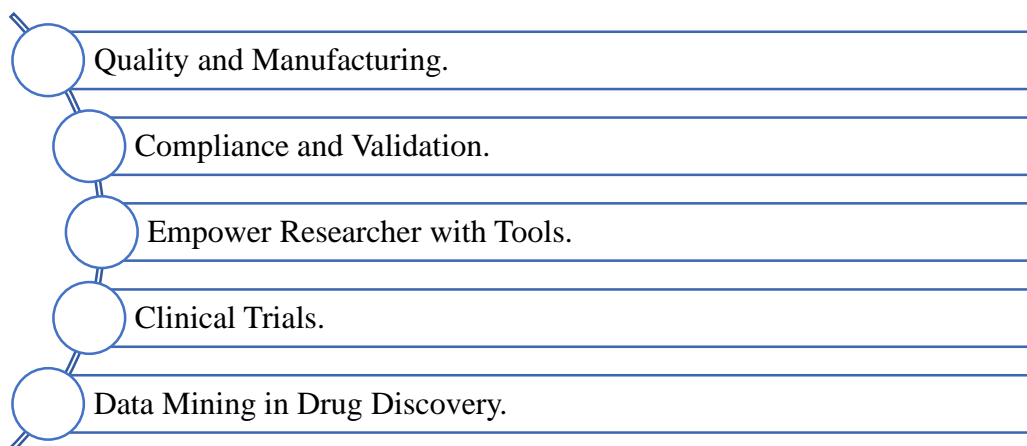
### 2.3.3 Type of analysis performed in SPSS

- I. Single sample *t*-test
- II. Two-sample *t*-test
- III. Paired *t*-test
- IV. Non-Parametric test for paired data
- V. Analysis of Variance
  - a. One way ANOVA
  - b. Two-way ANOVA

### 2.3.4 Applications of SPSS

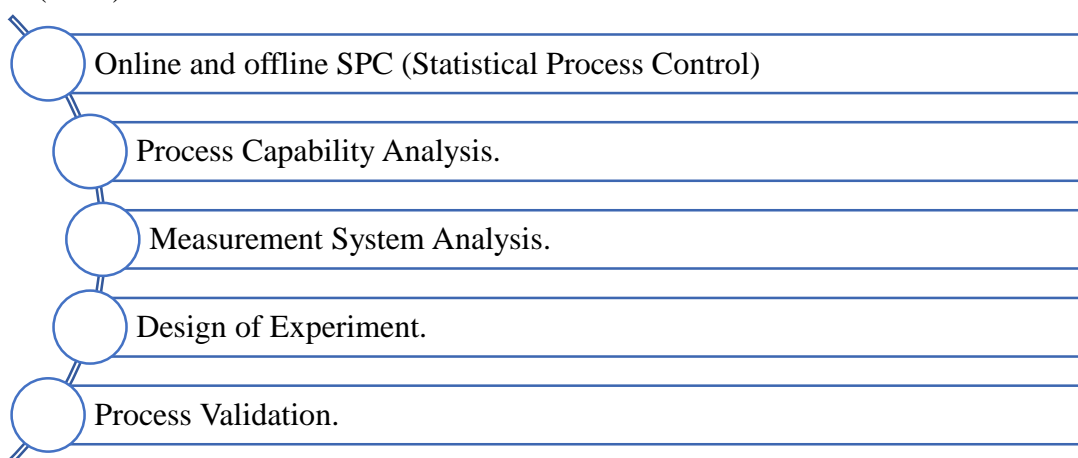
#### 1. Research and Development

SPSS given the technique of specialization in design, development, and implementation of corporate analytical and data mining solution integrated with the existing data acquisition and storage of systems. It continuously given the solution in pharmaceutical and medical research.



## 2. Quality and Manufacturing

SPSS can provide a single, standardized quality control and improvement solution that can be easily validated that cover the whole range of manufacturing statistics, including simple descriptive statistics and graphics, QC charting, capability analyses for normal and non-normal distributions, design of experiment (DOE) much more.



## 3. Compliance and Validation

SPSS offers a pre-validation core platform with fully secure user administration and configuration changes, data manipulation, analysis results and integrated data modules. Together these features can radically change the way of your organization works and remove the need for extensive manual documentation of analytical processes.

- i. Validation of Software.
- ii. Analysis of Audit Trails.
- iii. Computer System Validation.

## 4. Empower Researcher with Tools

Even small Improvement in research and development process directly affect an organization's bottom line many pharmaceuticals research organizations have experienced the positive impact of deploying SPSS.

Offering a choice of desktop (stand-alone), networks, and server based analytical platforms.

Statistician has provided analytical tools that are easy to use, relevant and integrated with their data source.

- i. Comprehensive Statistical Test & Graphical Tools.
- ii. Efficient use of limited Statistian & Training Resources.
- iii. Integrated with Existing System & Data Sources.
- iv. Rapid return on investment.

## 5. Regulatory Safety Testing

From simple bioassay and dose-response experiments to long-term survival and carcinogenicity studies. SPSS has the tools to handle your data. Where the standard study designed are used.

SPSS tools provide the most efficient mechanism for automating and validating routine analysis report.

- i. Power and Sample Size Calculation.
- ii. Predictive Data Mining.

- iii. Forecasting and Report Generation.
- iv. Analysis of Complex Drug's Assay.

## 6. Clinical Trials

Ethical consideration and at least in Phase I and Phase II Studies small sample size make it imperative that the most power statistical method be used in order to established efficacy and detect side effects. SPSS is unique in providing a fully document management system along with complete analytical toolkit. With ability to directly query standard database platforms and to import a wide range of data file format (including SAS files).

Generalized Linear and  
Non-Linear Model

Variance Component  
Estimations (ANOVA)

## 7. Data Mining in Drug Discovery

With the advent of new technologies that allow very large amounts of data to be captured and stored automatically many new challenges are presented.

In application of classical statistical method many were developed specially to extract the maximum amount of information from minimal data sets.

According to the FDA Critical Path Initiative the use of data mining technology and methodology is central to improvement discovery, preclinical and clinical research. SPSS perform an excellent crunch the complex data sets.

- 1) Handle Huge Data Sets with Ease.
- 2) Exploratory Data Mining Techniques.
- 3) Comprehensive Data Mining Toolkit.

### 2.4 GraphPad Prism (9)(10)(11)

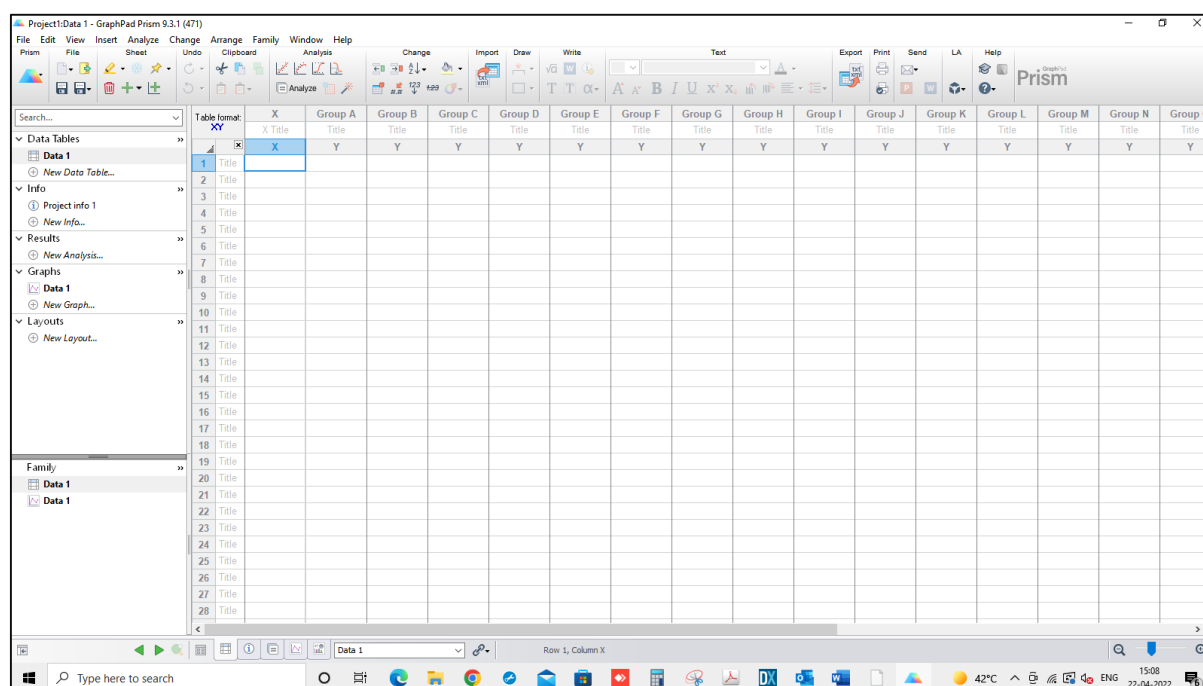
GraphPad Prism Statistical Analysis software focuses on scientific research as it allows scientists to analyze, graph, and present their scientific work with ease. Users do not need to have any programming expertise to analyze and create data visualizations with this software, as there is no coding required. Scientists can enter their data into tables structured for scientific research to guide them in statistical analyses and simplify their research workflow.

GraphPad Software was founded in 1989 by Dr. Harvey Motulsky, who at the time was on the faculty of the Dept. of Pharmacology at the University of California San Diego. In addition to doing research in receptor pharmacology, Dr. Motulsky also taught biostatistics to medical and graduate students and wrote the text Intuitive Biostatistics. Dr. Motulsky's pharmacology and statistics background make him uniquely qualified to provide researchers with the software features and support they require.

GraphPad Prism combines scientific graphing, comprehensive curve fitting (nonlinear regression), understandable statistics, and data organization. While it won't replace a heavy-duty statistics program, Prism lets you easily perform basic statistical tests commonly used by laboratory and clinical researchers. Prism offers t tests, nonparametric comparisons, one-, two- and three-way ANOVA, analysis of contingency tables, and survival analysis. Analysis choices are presented in clear language that avoids unnecessary statistical jargon.



The latest Prism version is **9.3.1 (Windows and Mac)**.



**figure: 7 initial screen of graphpad prism**

### 2.4.1 Key Features of GraphPad Prism (12)

**Statistical Analysis:** Prism offers a comprehensive library of statistical analyses, including nonlinear regression, t-tests, nonparametric comparisons, ANOVA, analysis of contingency tables, survival analysis and more. Users can confirm that they have chosen the appropriate test for their needs with a checklist that helps them understand the required statistical assumptions.

**One-Click Regression Analysis:** Prism allows users to select an equation and it will take care of the rest, including fitting the curve, displaying a table of results and function parameters, drawing the curve and interpolating unknown values.

**Data Tables:** Unlike spreadsheets or many other graphing programs, Prism has eight types of data tables that are specifically formatted for certain analyses, making it easier for users to enter data correctly and perform accurate analyses.

**Real-Time Updates:** When any changes are made to the data and analyses, those changes are reflected in the results, graphs and layouts instantaneously with updates in real time.

**Customizable Data Visualization:** Users can create a wide variety of visual representations of their data, including violin plots, box-whisker, bar graphs, sub column graphs, smoothing splines, scatter plots and more. They can also enhance these graphics with customization options that help them tell their data's story in whatever way they desire; they can choose the type of graph, how the data is arranged, the style of the data points, labels, colors, fonts, look and more.

**Prism Magic:** Users can select one or more Prism graphs and apply a consistent look to the set with one-click simplicity, saving valuable time spent otherwise on standardizing the look of multiple graphics.

**Export Graphics:** Users can export their graphs in high-quality and customize the file type, resolution, transparency, dimensions, color, space, etc. of their visualizations to meet the requirements of publication. To save time in the future, users can set their default export preferences.

**Work Automation:** Users can reduce the number of tedious steps needed to analyze data by setting up reproducible workflows that can create templates, duplicate families or clone graphs.

**Tools for Teamwork:** Prism allows for enhanced collaboration with other scientists, with all the information in a Prism project, including raw data, analyses, results, graphs and layouts, contained in a single, shareable file. Others can follow along with your work at every step, adding their insight and streamlining your collective research efforts.

### 2.4.2 Limitations of GraphPad Prism (12)

Some of the product's limitations include:

- 1) It is not available on Linux Operating System.
- 2) It is not available as a cloud-based platform.
- 3) It does not provide tools to write equations on a graph or layout. Users must instead find workarounds to insert their own equations.
- 4) It cannot create models with two or more independent X variables, which means it cannot calculate multiple nonlinear regression.
- 5) It cannot fit a model defined by a set of differential equations

### 2.4.3 Statistical Analysis in PRISM (13)

#### 1. XY data tables

- a. Linear regression
- b. Nonlinear regression
- c. Co-relations
- d. Analysis of covariance (ACNOVA)

#### 2. Column data tables

- a. One-sample, Two-sample and Paired *t*-test
- b. One-way analysis of variance (ANOVA)
- c. Nonparametric tests

#### 3. Grouped data tables

- a. Two-way ANOVA
- b. Repeated measures ANOVA

#### 4. Contingency data tables

- a. Chi-square tests
- b. Relative Risks, odds ratios, sensitivity and specificity

#### 5. Survival data tables

- a. Kaplan-Meier tests
- b. Long-rank tests

### CONCLUSION

A necessary component of surveys and studies is statistical software. The main pharmaceutical industries are actively embroiled in data-based learning systems for the development of pharmaceutical products. They are using a number of statistical tools to investigate more effective ways of manufacturing therapeutic products. There are multiple software tools available, and each one provides the user with slightly additional advantages. The software tool you choose will rely on a number of criteria, including your research issue, your statistical understanding, and your coding experience. In various circumstances, they can also reduce or eliminate a variety of biases and improve the process of developing medical products. The tools are frequently used in quality control and quality management. Process optimization takes these uncontrolled or noisy components into account and comprehends how they interact in dynamic processes. These elements might indicate that your data analysis is state-of-the-art, but like with all research, the effectiveness of the data you acquire depends on how well the study was carried out.

### CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

### REFERENCES

1. Eastwood BJ. Applied Statistics in the Pharmaceutical Industry With Case Studies Using S-PLUS. J Am Stat Assoc. 2003;98(461).
2. Top 10 Statistical Tools Used in Medical Research - The Kolabtree Blog [Internet]. [cited 2023 Mar 20]. Available from: <https://www.kolabtree.com/blog/top-10-statistical-tools-used-in-medical-research/>
3. Kenett R, Zacks S, Amberti D. Modern industrial statistics. With applications in R, MINITAB and JMP. 2nd revised and updated ed. 2013 Dec 27;
4. Stat-Ease [Internet]. [cited 2023 Mar 20]. Available from: <https://www.statease.com/about-us/history/>
5. Stat-Ease [Internet]. [cited 2023 Mar 20]. Available from: <https://www.statease.com/webinars/>
6. Stat-Ease [Internet]. [cited 2023 Mar 20]. Available from: <https://www.statease.com/about-us/newsroom/>

7. Stat-Ease [Internet]. [cited 2023 Mar 20]. Available from:  
<https://www.statease.com/publications/stat-teaser-newsletter/statteaser-0416/>
8. Salcedo J. Spss statistics for dummies, 4th edition: br. 2020 [cited 2023 Mar 20]; Available from:  
<https://www.wiley.com/en-us/SPSS+Statistics+For+Dummies%2C+4th+Edition-p-9781119560838>
9. Prism 9.4.0 Release Notes [Internet]. [cited 2023 Mar 20]. Available from:  
<https://www.graphpad.com/updates/prism-940-release-notes>
10. Company - GraphPad [Internet]. [cited 2023 Mar 20]. Available from:  
<https://www.graphpad.com/company>
11. GraphPad Prism - University IT [Internet]. [cited 2023 Mar 20]. Available from:  
<https://tech.rochester.edu/software/graphpad-prism/>
12. GraphPad Prism 9 User Guide - Installing Prism [Internet]. [cited 2023 Mar 20]. Available from:  
[https://www.graphpad.com/guides/prism/latest/user-guide/installing\\_prism.htm](https://www.graphpad.com/guides/prism/latest/user-guide/installing_prism.htm)
13. Ghoojani A. Advanced Statistical Methods and Applications (GraphPad Prism, SPSS, Minitab). 2019.