Graphing software for medical writers

Jesslynn Ooi[1,2], Daniela Traini[1,2], Paul M Young[1,2]

[1]*Respiratory Technology, Woolcock Institute of Medical Research, Sydney, Australia*
[2]*Discipline of Pharmacology, School of Medicine, University of Sydney, Sydney, Australia*

Correspondence to:

Paul M Young
Woolcock Institute of Medical Research, 431 Glebe Point Road, Sydney
NSW 2037, Australia
paul.young@sydney.edu.au

Abstract

The production of high-quality graphical figures of scientific data is an important aspect of medical writing, since the significance of the research findings can be lost to the audience if the data are presented poorly or inaccurately. This review highlights the various requirements that a writer may encounter when preparing data in a graphical format for presentation and gives an overview of 15 graphing software packages that are available on a number of platforms.

Keywords: Graphing software, Windows, Apple OS X, Linux, Figure preparation, Data analysis

Introduction

High-quality graphical figures of data should always be a priority when preparing a manuscript or presentation. The significance of a study or clinical trial can be lost to an audience if the data are presented poorly or inaccurately. It is important to choose a suitable graph format and represent your data in the clearest way possible. There are countless ways of representing datasets; the nature of your graph and the choice of graphing software will be dependent upon the type of data and the field of research. Depending on your data, you may wish to produce a scatter plot, a bar graph, a pie chart, or a multivariable three-dimensional (3D) plot. The type of graphical software that you settle on will be dependent on the output that you wish to generate.

When choosing a package it is important to consider the usability, the output formats, the programmability, and the availability of advanced features. Fifteen graphical programs across the Windows®, Mac OS X®, and Linux® platforms are briefly described and discussed in this review along with the basic requirements that a medical writer may have.

Identifying specific graphing requirements

The first step to choosing a suitable graphing software package is the successful identification of the user and the final output requirements. Medical writers may use multiple packages since not all the packages offer the same functionality, output options, and user interface. For example, a medical writer may require a graphical software package that has advanced statistical and regression features for one application, and an interactive 3D graph that can be outputted as a movie for another.

Generally, the best place to start when choosing a graphing package is to consider the following three questions:

1. How do you want to process and handle data?
2. What output format do you require?
3. What user interface are you comfortable with and what operating system are you using?

When answering these three questions, specific requirements should be considered. These requirements include: command line interface or graphical user interface, programmable functions, advanced regression features, output in multiple pixel and vector image formats, the ability to update graphs in real time, video output, and the inclusion of statistical functions and integrated graphing.

By asking these three specific questions, a medical writer will be in a position to narrow his/her choices to a few packages and then settle on the one that is most suitable.

Fifteen current scientific graphing packages

There are numerous graphing packages available to the medical writer. Here, we discuss 15 such packages across the Windows, Mac OS X, and

DOI: 10.1179/2047480613Z.000000000185

Linux platforms and highlight some of the key features that these packages offer.

DataGraph (Visual Data Tools)

DataGraph is a native Mac OS X software which is primarily used for clean 2D figure production. The list of graph types is restricted to basic scatter and line plots, histograms, basic function fitting (linear, polynomial, power, exponential, and arbitrary functions), bar charts, and scatter and line plots. This program deals with the data in columns and rows instead of individual spreadsheet cells, which indirectly increases the maximum handling capacity (it can deal with 'millions' of rows and 'thousands' of columns). The user interface for DataGraph is a series of panels that control each aspect of the output graph, allowing for very precise and easy adjustment of all the presentation aspects. This program's strength lies in its ease of use and quality figure production.

DataTank (Visual Data Tools)

DataTank is the big brother of DataGraph. It is capable of plotting 2D and 3D graphs from the templates of more than 80 graphs. Mathematically, it can handle 2000 computational actions and can compile 3D videos in QuickTime Virtual Reality format. DataTank is a native MAC OS X program and uses the Quartz and the Open Graphics Library (Open GL) programming languages to draw 3D graphics.

DeltaGraph (Red Rock Software)

Available on both the Microsoft and the Mac OS X platforms, DeltaGraph has an easy to use interface with Pantone® colour matching (Pantone being a proprietary 'colour space' used in the printing industry for standardised colour matching) and various raster and vector output options. The users can select from 25 built-in mathematical functions, 50 statistical functions, and 80 graphing templates (which include both 2D and basic 3D graph outputs).

Excel® (Microsoft)

Excel is a part of Microsoft's mainstay Office suite and many people use it for graphing purposes. While it is excellent for handling simple datasets, conducting routine calculations, and plotting data, it has a lack of graph types, and formatting and output options. Generally, medical writers will use Excel for handling and organising data before copying/exporting data into a scientific graphing program for figure production.

Grapher (Golden Software)

Grapher contains more than 60 types of 2D and 3D graphs covering features such as parametric function plots, basic function plots (x, y; and x and y as a function of z), polar graphs, ternary graphs, 2D and 3D vector plots, contouring, and basic fitting with 16 pre-determined fitting functions. The plots can be exported in a variety of vector- and raster-based formats. Grapher can also be automated by using the programming languages Visual Basic, C++, and Perl. This program is available only on the Windows platform.

GraphPad Prism (GraphPad Software)

While it was first designed for experimental biologists, this program is excellent at handling non-linear regression, curve fitting, and basic statistical tests (t-tests, non-parametric comparisons, one- and two-way analysis of variance (ANOVA), analysis of contingency tables, and survival analysis). It is very popular and allows for quick no-fuss statistical calculations, which integrate directly into the outputted graphical figures. GraphPad Prism is available on both the Microsoft and the Mac OS X platforms.

IGOR Pro (WaveMetrics)

The target market for IGOR Pro, which is capable of producing quality 2D and 3D plots, is the science and engineering disciplines. Curve analysis is the mainstay of this program and it can handle curve fitting, peak analysis, signal processing, and basic statistics with ease. Of particular interest to medical writers is the 3D plotting program, which is useful for the analysis and the presentation of 3D volumetric data, e.g. from medical magnetic resonance imaging (MRI) or X-ray computed tomography (CT) scans. It is possible to create short 3D animations through manual manipulation or through programming. Moreover, the users may install third-party plugins (or write their own in C or C++) to acquire data in real time. IGOR Pro is available on both the Microsoft and the Mac OS X platforms.

KaleidaGraph (Synergy Software)

KaleidaGraph is a 2D scientific graphing software package intended for statistical analysis and graphics production. It has a range of 2D graphs, including line and scatter plots, probability graphs, histograms, and stacked, column, polar, pie, and function charts. This program allows for fine control over presentation, and boasts of extensive curve fitting functions and features such as the ability to plot up to 99 dependent variables on a

single set of axes. The formulae are available in the datasheet and standard templates can be constructed. The program can run on both the Windows and the Mac OS X platforms.

MagicPlot (Magicplot Systems)

MagicPlot is a cross-platform graphing software based on Java for plotting 2D *X*–*Y* graphs in the Windows, Mac OS X, and Linux systems. Most functions are basic but the program does include nonlinear curve fitting and 2D waterfall graphing functions.

Maple (Maplesoft)

Maple's targeted application areas are financial modelling, statistics and process control, physics, and control design. The statistics package can analyse distribution parameters, compute cross-correlation and autocorrelation, smooth data to various degrees, and perform 11 types of hypothesis testing. Aside from the statistics, Maple is able to generate 170 plot types, including 3D, vector field, contour, polar, conformal, and density plots. In addition, Maple can conduct ordinary and partial differential equations and produce statistical plots. The program allows the user to fine-tune the graphical output (e.g. surface style, lighting, and glossiness) of 3D plots. Maple also features a code generation option, which allows users interested in programming to develop a prototype or solution in the Maple language and to generate Visual Basic, MatLab, Java, C, C#, or Fortran code. Maple is available across the Windows, Mac OS X, and Linux platforms.

Mathematica (Wolfram Research)

Mathematica is a high level and extremely diverse program suitable for multiple applications, including pure statistical and data analysis, social network analysis, 3D volumetric and imaging data, control systems, signal processing, engineering, parallel computing, wavelet analysis, financial engineering, geographical analysis, and computational biology. It claims to have the 'largest collection of algorithms in a single system', which is relevant for users who wish to perform enhanced statistical analysis. The program is also capable of creating and analysing volumetric and 3D imaging data, which is useful for individuals who wish to create 3D slices or pattern recognition algorithms from MRI or X-ray CT scans. Mathematica also has a suite of highly customisable bioinformatics functions, including computable genome and protein data analysis, sequence alignment, and protein structure rendering. While these features are not specifically graph related, they may be of use to medical writers who need to present genome- or protein-related data in a publishable form. Mathematica is available across the Windows, Mac OS X, and Linux operating systems.

MATLAB® (MathWorks)

MATLAB is a high-level language intended for numerical computation, programming, graphing, and application development. It has a wide range of science and engineering applications, including signal processing and communications, image and video processing, and computational biology. The program is capable of applying numerical computation methods to analyse data and develop algorithms, and can be used to output both 2D and 3D models. It can easily deal with linear algebra, statistics, Fourier analysis, filtering, optimisation, integration, and differential equations. Data acquisition in real time is possible with add-on products. MATLAB is touted on its website to be 'faster than traditional software' when it comes to data acquisition, visualisation, and analysis. New users should allow some time to familiarise themselves with the MATLAB language. The program is available for the Windows, Mac OS X, and Linux systems.

Numbers (Apple Inc.)

While visually appealing and functionally easy to use, this program is not targeted towards the scientific community. However, for the medical writer who wishes to produce sharp presentation graphics for media purposes, Numbers has a range of 3D graphics options with realistic finishes (i.e. bar graphs or pie charts with polished metal, cloth, or wood grain textures). In addition, basic curve fitting and error bars can be incorporated into some 2D graphs. Recently, Numbers has been updated to integrate between Apple's desktop OSX and mobile iOS platforms via iCloud. This makes Numbers a feasible choice for medical writers who are 'on the go' and require a tablet option that synchs seamlessly with their desktop version at home.

OriginPro (OriginLab)

OriginPro is predominantly a data analysis software package capable of polynomial, linear, and nonlinear curve fitting, peak analysis (subtraction, detection, integration, fitting, batch analysis), statistics (descriptive statistics, hypothesis testing, ANOVA, non-parametric analysis, multivariate analysis, survival tests, power and sample size tests, receiver operating characteristic curves), and

mathematical function analysis (interpolation, extraction, differentiation, integration, area and volume calculation). Newer versions of OriginPro have introduced 3D graphing functions (3D parametric plots, waterfall functions) and 3D animations. OriginPro is a Microsoft-based program.

SciDAVis (Open source)

A free interactive application with similar functions to OriginPro and SigmaPlot (see below), SciDAVis aims to deliver publication-quality 2D and 3D plots. Cells in tables can be computed to standard level and special functions if coupled with the Python programing language. Analysis of data and curves includes multiple curve fitting, filtering, interpolation, non-linear curve fitting, fast Fourier transform analysis, correlation, and the deconvolution of data. Plots can be exported in a variety of vector- and raster-based formats. This software is available across the Windows, Mac OS X, and Linux operating systems.

SigmaPlot (Systat Software)

SigmaPlot, a Windows-based software package, has a number of functions relevant to scientific research and is operated through its easy to use graphical user interface. SigmaPlot 12.5 (the latest version at the time of writing) boasts of more than 100 2D and 3D graph templates, including 2D vector plots, 3D mesh plots, contour projections, radar plots, and dot density graphs. One of the features of this program is the ability to embed graphs in Microsoft Word® or Excel through the Visual Basic script; medical writers who prefer to operate on PC platforms may find this useful. In addition, tasks can be automated within SigmaPlot in macro language. Descriptive statistical analysis can be performed easily alongside more advanced tests such as Deming regression analysis, three-way ANOVA, and non-parametric one-sample *t*-tests. SigmaPlot is also capable of data transformation. An interesting aspect of the SigmaPlot program is the inclusion of an Enzyme Kinetics Module, which comes with additional analysis features and graph types (e.g. Lineweaver–Burk, Eadie–Hofstee, and Scatchard graph types).

Summary

There are many graphing packages available to the medical writing community and the choice of package will be specific to the dataset and the output requirements of a particular project. Many of the packages have simple to use graphical interfaces, while others have command line interface options and integration of a number of programming languages. The output formats also vary between packages, although the majority now offer common graphical formats required by most journals. Ultimately, the best way to identify the most suitable package is via hands-on experience. Luckily, most packages offer users the opportunity to 'try before you buy' and thus, after identifying a few potential candidates, the user can process their data by using a few options before settling on their final choice.

Acknowledgements

PMY is the recipient of an Australian Research Council Future Fellowship (project number FT110100996). DT is the recipient of an Australian Research Council Future Fellowship (project number FT12010063). JO is the recipient of an Endeavor Scholarship.

Author information

Jesslynn Ooi is a PhD student at the University of Sydney and the recipient of multiple early career awards, including a European Framework 7 scholarship in nanotechnology and an Endeavour Research Fellowship.

Daniela Traini is an Australian Council Future Fellow and Associate Professor at Sydney Medical School.

Paul M Young is Head of Respiratory Technology at the Woolcock Institute of Medical Research and Associate Professor at the Sydney Medical School, University of Sydney.

The authors are strong supporters of early career researchers and founding members of http://www.ecr2star.org.