- [The Best Graphing Software](). Solving the problems with Office.
- Updates: [Sample size](); [Clinical inferences](); [Validity and reliability]()

[Reprint pdf]() · [Reprint doc]()

The Best Graphing Software

Will G Hopkins, Sport and Recreation, AUT University, Auckland, New Zealand. **Email**. Sportscience 15, i-iii, 2011 (sportsci.org/2011/inbrief.htm#graphs). Published June 2011. ©2011

Update Aug 2018. Bill Microsoft Gates has at last solved the problem (detailed [below]()) of lines turning into tram tracks when graphics from Excel are pasted into Powerpoint as a metafile and ungrouped for further editing. Thanks, Bill, you got there in the end! The latest version of Powerpoint in Office 2016 also has a wonderful new feature for lining up objects, whereby dashed lines appear when edges of an object you are moving line up with edges of other nearby objects. Now, Bill, please reinstate Recent Folders in the Save As… windows in Windows 10.

Update Feb 2018. I still use Excel to create graphs, and I avoid the thick lines generally by not taking apart the graph after pasting into Powerpoint. Instead, I clean up the graph by covering items with white boxes and adding new numbers, text and lines. I would use GraphPad Prism instead of Excel, if I had it. Graphs in the new free SAS Studio package can be taken apart in Powerpoint, but they require too much programming for the casual user. See examples in the [suite of materials on mixed modeling in SAS Studio]().

In April this year (2011) I posted [a message]() to the [Sportscience mailing list](), asking for advice on software for creating publication-quality graphs. I was hoping to solve a problem I had encountered with the use of the 2007 Office suite for graphing, as described in [an earlier in-brief item](). Here is a restatement of my problem and the various solutions, based on the dozen or so responses I received and further experimentation.

The Problem: Bugs in Office 2007/2010

I have been using Excel 2003 to create graphs, which I then paste into Powerpoint 2003 to take apart and clean up for slideshows and/or manuscripts. My problem is that the Excel+Powerpoint combo in Office 2007 and 2010 has unacceptable bugs for such processing, and my IT people are telling me I have to give up Excel and Powerpoint 2003 when they install Windows 7 on my laptop soon. So I am looking for graphing software that will make graphs I can modify with the latest Powerpoint. In any case, I am interested in software that does a better job than Excel, although I still have to use Excel: all my spreadsheet resources are in Excel, so anyone using the spreadsheets, including me, will still want to transfer figures from Excel into Powerpoint.

Solution #1: Keep Using Office 2003

I have Jim Martin to thank for sharing similar problems and strongly advising me to keep Office 2003 going. He said I could buy Office 2003 for US$130, if need be, "a pretty inexpensive solution". I agree. I will need technical help to get 2003 and 2007/10 running together, as I did for my current setup with Windows XP. I will hire an outside geek if my IT people refuse to help.

So here's what to do when you want to use the graphs in Excel to make graphs that look good in Powerpoint presentations and that are acceptable in publications (adapted from my [in-brief item]() on preparing graphics for publication): open the xls or xlsx spreadsheet with Excel 2003; copy the graph to the clipboard; paste-special as a Windows metafile; click-drag to change the size; right-click and ungroup twice; edit the graph elements (easier if you first delete the white background that comes through with the graph); resize again if necessary (by ungrouping all grouped items, cutting to the clipboard, re-pasting as a metafile, re-sizing...); finally export it for publication, or copy the graph or the whole slide into Powerpoint 2007 or save the whole presentation as a pptx for presentation.

The only alternative to using Powerpoint

2003 is to edit the Excel graphs in a vector drawing editor before pasting into Powerpoint 2007. I searched for freeware editors and tried Inkscape, but it doesn't take Excel graphs apart, and I don't know whether expensive high-end editors like CorelDraw and Macromedia Freehand will do the job. In any case, you can regard Powerpoint 2003 as a cheap-ware powerful drawing editor that has a great additional advantage: the ability to resize graphics and keep everything, including font sizes, in proportion, as described above.

By the way, one of the only advantages of Powerpoint 2007 over 2003 is the ability to edit an element of a group without ungrouping (which in 2003 results in loss of the animation for the group). I sometimes go into Powerpoint 2007 solely for this operation, then scuttle back to 2003. It's bizarre that the Microsoft developers could have an eye for this kind of detail, yet ignore the graphing needs of the scientific community.

Solution #2: GraphPad Prism

Many of you (Stephen Seiler, Stuart Galloway, Felipe Carpes, Catherine Bacon, Conrad Earnest, Paul Montgomery) sang the praises of Prism and pointed out that its graphs come apart in Powerpoint. I downloaded a trial version to see, and yes, it's true. In Powerpoint 2003 the formatting is retained perfectly, whereas in Powerpoint 2007 lines lose the thickness you have given them in Prism, but that's not a major problem. I also checked out Prism's statistical capabilities: it's more powerful than Excel but not as powerful as Statistica. A problem with GraphPad is cost: around $US450 for a single academic license. I won't be buying one yet.

Prism will open an Excel spreadsheet of data, but any graphs in the spreadsheet get corrupted. Prism is therefore an adjunct for Excel, not a replacement.

Solution #3: SigmaPlot

Several people (Martin Buchheit, Mike Hamlin, Yngve Sommervoll) are happy with the graphing and other capabilities of SigmaPlot. I got conflicting advice about pasting into Powerpoint, even from the SigmaPlot site, so I downloaded a trial version, created a graph, pasted it into Powerpoint as a Windows metafile, then took it apart without any problems. Price: US$550.

Solutions #4 and more

The open-source R package is renowned for its graphics. It's also free, incredibly powerful, but user-toxic when I last tried it a couple of years ago. A Windows interface is available and may help newbies like me get started. Rob Rein, who runs the R tutorial group, sent me an R graph as a Windows metafile, and it came apart perfectly in Powerpoint 2007.

Dan Heil was ecstatic about graphs in Statistica. I checked with our local Statistica agent, who sent me a graph that edited perfectly in Powerpoint 2007. Price: ~US$675 (I was quoted NZ$850+tax). Statistica probably represents better value than SigmaPlot. Statistica even has a mixed model, but it isn't good enough yet for our purposes. I periodically ask them to improve it.

Michael Brach: "I'm working with Linux while most of my colleagues use Windows, I'm always looking for software which uses open formats and/or is available for both operating systems. Thus, we use DIA, for diagrams not based on numbers (e.g., flow charts) and GNUPLOT, for making graphs from statistical and other data. Gnuplot can also very easily be used in batch mode, thus producing lots of graphs reading a list of variables, a list of subjects, and so on. Both applications are free and can save graphics in formats that Powerpoint can read."

Filipe Carpe recommended Microcal Origin. He said Powerpoint will edit its graphs. The Pro version is US$800.

I haven't evaluated Stata, but from what I could find at its website, the graphs are editable in Powerpoint. Price depends on the version you want: US$600 to $1600+.

Other Solutions and Non-solutions

I had hopes for the free OpenOffice suite, which features a spreadsheet very similar to Excel 2003. It opens my spreadsheets, but the graph symbols do not have outlines and, incredibly, there is no plain circle symbol. The OpenOffice equivalent of Powerpoint is quite klunky and too different from Powerpoint to be worth persevering with. There were also too many bugs when transferring a slide to Powerpoint 2003 or 2007.

Conrad Earnest suggested Kalidegraph. It looks promising, because it's not too expensive (US$140), but I searched the site without success for anything about ability to edit in Power-

point. Conrad also pointed me to a comprehensive [list of graph software](#) at Wikipedia. Kevin Short pointed out that Kalidegraph works on PC and Mac, "though it's still not easy to show a student quickly how to use it and copying images to other formats like PPT results in loss of image quality in our hands."

James Best used [qtiplot](#) ("very similar to Origin graphing software") to generate a graph, which he turned into a file (using a [converter program](#)) that Powerpoint could open. I was able to ungroup it to produce editable lines, but text became un-editable images.

Ken Quarrie: "You might want to check out [SCIDAVIS](#) as a tool for plotting/charting and putting into documents. It is free for download. It can read in ascii files and has quite an array of functions." I checked out the site and discovered that it is an offshoot from qtiplot. I was not encouraged when I found that the site was last updated more than a year ago.

Alan Batterham suggested [XL toolbox](#), a free add-in for Excel that allows assembly of several graphs into a single panel of graphs for publication. Unfortunately he found that it does not allow editing of the graphs.

Finally, SAS and SPSS produce awful graphs that can't even be taken apart–a disgrace for such high-end expensive statistics packages. It takes years to learn SAS, but it's still worth the effort, in spite of the graph problems: you can read in data from the messiest text files, and it has the most wonderfully powerful state-of-the-art analytical procedures. I routinely use Excel to create graphs from the text output of SAS. SPSS is limited to data only in the usual column format. On the analysis side it now has a mixed model that will handle complex repeated measurements, but you have to use dummy variables to make it work properly–instructive for your understanding of linear models, but not recommended for routine use. See [Tips and tricks for SPSS](#) if you have no choice.

Updates: Sample Size, Clinical Inferences, Validity and Reliability

Will G Hopkins, Sport and Recreation, AUT University, Auckland, New Zealand. **[Email](#)**. Sportscience 15, iii, 2011 (sportsci.org/2011/inbrief.htm#updates). Reviewer: Alan M Batterham, University of Teesside, Middlesbrough, UK. Published June 2011. [©2011](#)

Sample Size. Updates for the [spreadsheet](#) are detailed at the beginning of the [article](#) on sample size. In summary, there is now a panel for sample size when the dependent variables is a count (e.g., of injuries or of events in a game), and the panel for single event outcomes now allows inclusion of smallest beneficial and harmful effects as risk difference, odds ratio and hazard ratio (in addition to the risk ratio that was there originally). The article also has a bullet point for sample size in a pilot reliability study for a controlled trial.

Clinical Inferences are now included in the [controlled-trial spreadsheets](#) (pre-post parallel-groups, post-only crossover, pre-post crossover) and in the [spreadsheet for comparing group means](#). The changes represent implementation of the principles explained in [the article](#) on clinical, practical and mechanistic inferences. Briefly, a clinical inference is based on the notion that a treatment or other effect should be used on people only if the risk of harm is sufficiently low (<0.5%, "most unlikely") and the chance of benefit is sufficiently high (>25%, "possible"). The spreadsheet allows these thresholds to be made smaller (more conserva-tive) via simply increasing the level for the confidence interval from the default of 90% that is used for mechanistic inferences, which thereby also become more conservative. There is also provision for controlling the error rate when there are several independent inferences, simply by inserting the number of inferences.

I have included cells showing a less conservative approach based on the odds ratio of benefit to harm, whereby an effect can be used if odds of benefit outweigh the odds of harm by more than 66. (See [the article](#) on clinical inferences for more.) Reducing the threshold probabilities for harm and benefit does not increase this odds ratio, and I have been unable to devise a systematic approach to increasing the odds ratio for more conservative inferences. The odds ratio should therefore be used sparingly or maybe only once in a study.

Some cells in the spreadsheets have extensive comments to guide their use. In particular, there is an explanation about adding up chances of harm and benefit with multiple inferences as a more efficient alternative to reducing the thresholds of 0.5% and 25%. However, I sus-

pect this approach will require familiarity with magnitude-based inferences that will take researchers months or years to acquire, if they bother at all. It's much easier to stay with null-hypothesis testing and drop the alpha level, even if it means you end up with "no effects"!

This update has been a long time coming, because the changes were complex and time consuming. If you spot any minor errors or formatting issues, please get back to me.

————

Validity and Reliability. I dealt with these measurement concepts in an article/slideshow devoted to assessing athletes published here in 2004, but I have been using a more generic updated version for teaching these concepts to graduate students. The slideshow has a reference list of various relevant articles, amongst which is something you shouldn't miss: the Socratic dialogue on comparison of measures (i.e., validity studies) published here in 2010.