Statistics
in Medicine

Special Issue Paper

Received 27 November 2013, Accepted 17 May 2015 Published online 25 June 2015 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.6549

Seeing is believing: good graphic design principles for medical research

Susan P. Duke,[a*†] Fabrice Bancken,[b] Brenda Crowe,[c] Mat Soukup,[d] Taxiarchis Botsis[e] and Richard Forshee[e]

Have you noticed when you browse a book, journal, study report, or product label how your eye is drawn to figures more than to words and tables? Statistical graphs are powerful ways to transparently and succinctly communicate the key points of medical research. Furthermore, the graphic design itself adds to the clarity of the messages in the data. The goal of this paper is to provide a mechanism for selecting the appropriate graph to thoughtfully construct quality deliverables using good graphic design principles. Examples are motivated by the efforts of a Safety Graphics Working Group that consisted of scientists from the pharmaceutical industry, Food and Drug Administration,[‡] and academic institutions. Copyright © 2015 John Wiley & Sons, Ltd.

Keywords: graphics; communication; visualization; visual perception

1. Introduction

In today's world of ever more (and more accessible) medical research data, quality visualizations are critical to appropriate interpretation of the data, particularly for those decision-makers (upper management, regulators, healthcare providers, and patients) with a non-statistical background [1]. Appropriate graphic design plays a critical role in the creation of graphs that efficiently and effectively translate the key clinical messages in the data. Many statisticians have described effective graphics, both decades ago and more recently [2–15]. Well-designed graphics have also been encouraged by pharmaceutical statisticians [16–25].

While many would agree that well-designed graphs are powerful communication vehicles, a brief look at medical research publications and regulatory submissions indicates there is room for improvement [18, 20–22]. This paper aims to explain more about effective (and ineffective) graphs. For a graph to be effective, it must be easy for its audience to decode and interpret. Because human brains are so good at pattern recognition [14], utilization of the most powerful techniques that render the key messages of the graph via patterns is especially important.

Tufte stated, 'of all methods for analyzing and communicating statistical information, well-designed graphics are usually the simplest and at the same time the most powerful' [10]. Cleveland stated further, 'When a graph is constructed, information is *encoded*. The *visual decoding* of this encoded information is *graphical perception*. The decoding is the vital link …no matter how ingenious the encoding …and no matter how technologically impressive the production, a graph is a failure if the visual decoding fails' [11].

We, the authors of this manuscript, were members of a Safety Graphics Working Group that included scientists from the pharmaceutical industry, Food and Drug Administration (FDA), and academic institutions. The goal of the working group was to create a publicly available repository of graphics designed

[a]*GlaxoSmithKline, Research Triangle Park, North Carolina, United States*
[b]*Novartis AG, Drug Safety and Epidemiology, Novartis Campus, Basel 4056, Switzerland*
[c]*Eli Lilly and Company, Global Statistical Sciences, Lilly Corporate Center, Indianapolis, IN, 46285, United States*
[d]*Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, Maryland, United States*
[e]*Center for Biologics Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, Maryland, United States*
[*]*Correspondence to: Susan P. Duke, GlaxoSmithKline, Research Triangle Park, North Carolina, United States.*
[†]*E-mail: susan.p.duke@gsk.com*
[‡]*The opinions and information in this article are those of the authors and do not represent the views and/or policies of the U.S. Food and Drug Administration.*

for reporting on clinical trials safety data, including detailed explanations and computer code for each graphic design. The work is published on a collaborative web platform (i.e., wiki) [26]. The graphs on the wiki are a helpful starting point for developing medical research graphics. We authors were also all members of the 'General Principles' subteam, which focused on identifying graphics principles and best practices. Our work is also available from the wiki and includes nine best practices for making graphs (see Appendix). The participation of FDA scientists in the working group should not associate this paper with an FDA effort or initiative to publish any guidance on visualization principles and best practices.

Bearing in mind that a graph is another type of statistical method, and being that figures are as compelling as they are, why *wouldn't* a statistician use this valuable method to communicate the key messages of the data to their audience? We believe that effective communication is a core responsibility of the statistician. Our motivation is to provide recommendations to visually present data from medical research that enhance the ability to effectively convey key medical research findings.

This paper covers several important topics. First, Section 2 contains comprehensive information that can be used to choose an optimal graph for a given dataset. It describes the influential factors and presents a flow diagram/decision tree for selecting the optimal graph type [27]. We refer to several graphs that will be discussed in detail in later sections to illustrate the general principles we are introducing in this section. Second, Section 3 contains examples of various useful graphs that are more complex than are often utilized in medical research. In particular, Section 3.3 may seem difficult to those who are unfamiliar with network visualizations, but we present it as a first glimpse into an important emerging area of visualization. Finally, Section 4 contains several examples that demonstrate application of the flowchart to real-world situations and illustrates how to improve graphs by applying best practices.

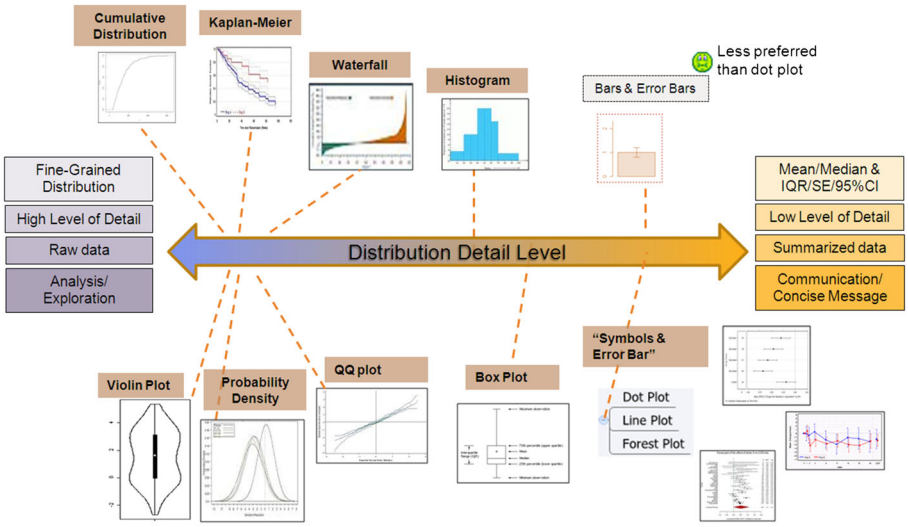
2. What graph is most appropriate?

One of the challenges in graphical representation of data is the choice of the optimal graph type for a given dataset. Numerous options exist for common safety questions/domains, as in Amit *et al.* [28] and further expanded by the Safety Graphics Working Group [26]. Other examples can also be taken as a starting point and potentially improved using best practice recommendations for graphics (see Appendix). But, the question of the optimal graph for other types of data remains. As an aide in choosing the optimal graph type (especially to answer novel questions), we describe the influential factors and then present a flow diagram for selecting the optimal graph type [29]. Later in this manuscript, we provide several graphical examples that we will reference in this section to illustrate the principles of choosing an optimal graph.

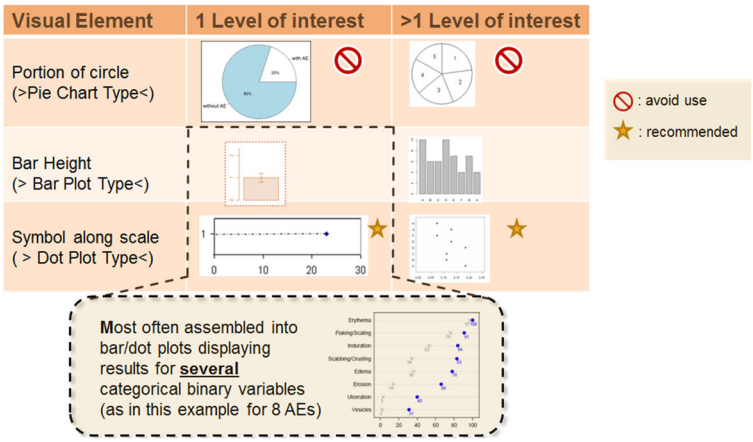
2.1. Choosing the graph type for one variable of interest

There are many ways to approach the choice of a graph. To identify the influential factors of such a choice, let us take the example of Figure 7(a). The underlying variable displayed is quantitative: systolic blood pressure. What is actually displayed in the first red bar of the figure is not the whole distribution of this variable for all the subjects in the study, but rather summary-level information, the mean and corresponding confidence interval (CI). This leads us to the first factor that influences the type of graph: the level of detail to convey from the distribution of the variable of interest. When reducing the distribution to its mean and 95% CI, a dot plot with error bars can be used as a graphical element to represent those quantities. However, if more details are to be provided from the distribution of the variable, one needs to use graphical elements conveying more characteristics of the underlying distribution (Figure 6(a) and b), such as the boxplot, a histogram, or the empirical probability or cumulative distribution function. Figure 1(a) summarizes graphical elements that are typically used for representing distributions of a continuous variable, by desired level of detail.

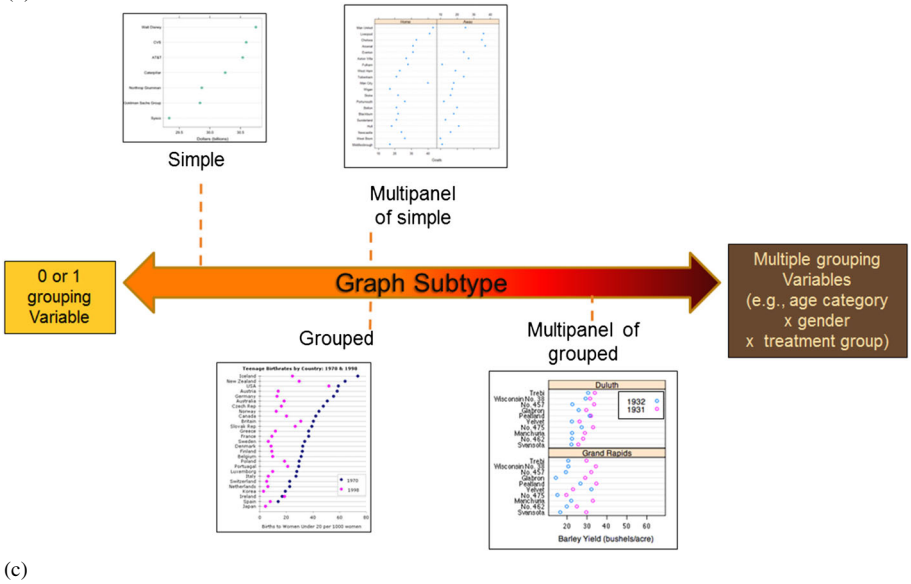
The level of detail of the distribution has less impact for categorical variables, where one would usually describe the probability of each level of the categorical variable. Let's take the example of Figure 10(a) and focus on the first graphical element starting from the left, the dark blue bar. The underlying variable is a binary one and is an indicator of whether or not a subject had any eye irritation during some time interval. The distribution of a binary variable is typically summarized by the probability of 'Yes' and the probability of 'No'. Because the sum of those is 1, one usually displays only one of these probabilities, and this is what the first bar represents, using as an estimate: the percentage of subjects treated with placebo who had eye irritation during week 1. When a single quantity (probability of 'Yes') needs to be displayed, a graphical element such as a dot or a bar can be used to represent that single quantity. For non-binary categorical variables, the probability of each level needs to be conveyed, which a set of bars or dots can be



(a)



(b)



(c)

Figure 1. (a) Graphical elements to summarize a quantitative variable's distribution, by desired level of detail. (b) Graphical elements to summarize a categorical variable's distribution as a function of the number of levels of the categorical variable. (c) Graph subtypes to represent different subsets of the data. IQR, interquartile range; SE, standard error.

used to represent. Alternatively, a pie chart or a stacked bar plot can be used (although not recommended as per best graphing principles), where a portion of the pie or a portion of a bar represents the probability of a level. Figure 1(b) summarizes graphical elements that are typically used for representing distributions of a categorical variable.

The approach to displaying the distribution is *influenced by the variable type* (quantitative or categorical) under consideration, and *the number of levels of involved categorical variables*. These considerations define the second influential factor in the choice of the graph.

Let us now go a step further. So far, we have seen how a graphical element may be used to represent the distribution of a quantitative or categorical variable. Taking a look at Figure 7(a) again, one can see that it is not just displaying one single graphical element but contains multiple bars, representing characteristics of the distribution of the variable of interest for different *subsets* of the data. In the case of Figure 7(b), it displays the mean and 95% CI for the systolic blood pressure *for groups of patients of different gender and having been exposed to different doses of a drug*.

The number of these subsets and how they are defined will influence the arrangement of the different elements on the graph. In particular, a significant role is played by the number of grouping variables (categorical or quantitative discrete) used to define those subsets and the number of their levels, because we need to provide the audience with ways to identify which graphical element corresponds to which subset. In Figure 7(a), the grouping variables are 'gender' (two levels) and 'treatment group'(four levels). In Figure 10(a), the grouping variables are 'treatment group' (three levels) and 'time point' (six levels). These examples show that there are several ways for the audience to identify the link between graphical elements and the subsets they represent:

- Varying the visual properties (symbol type, symbol color, and error bar color) of the graphical elements for the different levels of the 'treatment group' variable. This gets best results when there are no more than four levels and is often referred as a 'grouped' plot (Figure 10(b)). The relationship between the visual properties and the level of variables need to be conveyed. This is usually done through a legend. However, legends can be avoided by use of self-explanatory symbols or annotations, which usually make the information within the graph easier to decode.
- Using the other plot axis (when only one axis is used for the primary variable of interest) (e.g., Figure 10(b) for 'time point'). In the authors' opinion, this gets best results for quantitative discrete variables, when there are a large number of levels, and when the categorical variable is ordinal (levels can be displayed in increasing or decreasing order on the axis).
- Producing multiple plots (also called multipanel plots), with each panel representing different levels of the variable (e.g., Figure 6(a) for the 'time point' variable). This gets best results when the other options provide cluttered graphs.

As the number of subgrouping variables increases, a combination of those approaches should be used (e.g., grouped dot plots in Figure 7(b) or multipanel of grouped Kaplan–Meier plots in Figure 8). Multipanel plots are concepts introduced by Tufte [10] ('small multiples') and Cleveland [11] ('trellis plots'). The optimal display (and which variable should be represented by which approach) will often be obtained by a balance between bringing as close as possible elements in the graph that should be compared (grouped plots with legend) and avoiding cluttering (use of multiple panels/graphs).

In summary,

- For quantitative variables, the most important decision factor will be the level of distribution detail to display (Figure 1(a)), whereas for categorical variables, the number of its levels (Figure 1(b)).
- The choice of how to assemble (grouped plot, multipanel plot, etc.) graphical elements into the graph to represent different subsets of the data will be driven by the number of grouping variables defining these subsets as well as the number of levels they hold (Figure 1(c)).

2.2. Extension to bivariate distributions and higher-dimension distributions

While the approach described in the preceding text will cover a large majority of the graphs used in the clinical trial setting and beyond, it still relies on the fact that there is a single primary variable of interest, for example, 'systolic blood pressure' or 'presence of eye infection'. However, we are often interested in the association between two or more variables. For example, the maximum alanine aminotransferase versus the maximum total bilirubin for a subject while on-study. If both variables are quantitative, the level of detail of the multivariate distribution will determine what the graphical elements will be. This

can range from a low-detail mean and 95% CI ellipsoid to hexagonal bin (hexbin) plots, scatterplots, sunflower plots, or scatterplot matrix graphs for higher dimensions.

The same considerations as in the univariate case hold when it comes to assembling the graphical elements across different subsets of data: adopting a grouped plot, multipanel plot, or combination of those will be driven by the number of subsets and their corresponding categorical variables.

2.3. Beyond the graph selection algorithm

The nature of the data and the goals of the visualization can inform the choice of a type of graph, but the described selection algorithm may leave the user with more than a single possibility. In that case, analysts must consider all of the principles of statistical graphics listed in the Appendix and illustrated in Section 4 to rank those possibilities. In case of equal ranking, analysts will need to apply their own judgment about the best way to effectively communicate the meaning of the data. Feedback from your colleagues can be very helpful. Show your draft graphs to a few colleagues who have a perspective similar to the intended audience and ask them what they see in the draft version. The best graphs, like the best writing, require multiple iterations to hone the key messages.

3. More complex graphs

Several common graph types were discussed in the previous section. They usually are excellent choices when presenting small data. However, some of them may be problematic for datasets with a large amount of information, for example, when plotting patient-level information – a common attribute of later-stage clinical trials. Multiple line graphs, commonly called spaghetti plots, and scatterplots are two graph types that may suffer from overplotting of data points.

It is often necessary to illustrate complex relationships in big data involving, for example, exposures to combinations of drugs and/or biologics and multiple, related adverse events (AEs). Network visualizations may support this task. The data are visualized as network structures with the key entities and their relationships represented as the nodes and the edges of the networks, respectively. Network visualizations are particularly useful for qualitative evaluation of data, and they may suggest further analysis of the complex relationships between the key entities in large datasets.

3.1. Spaghetti or lasagna?

Clinical trials often have repeated measures for some outcome for the subjects enrolled in the study. Line graphs are commonly used to display repeated-measures data, and a collection of line graphs for multiple subjects is often referred to as a spaghetti plot because the overlapping, irregular lines may resemble the

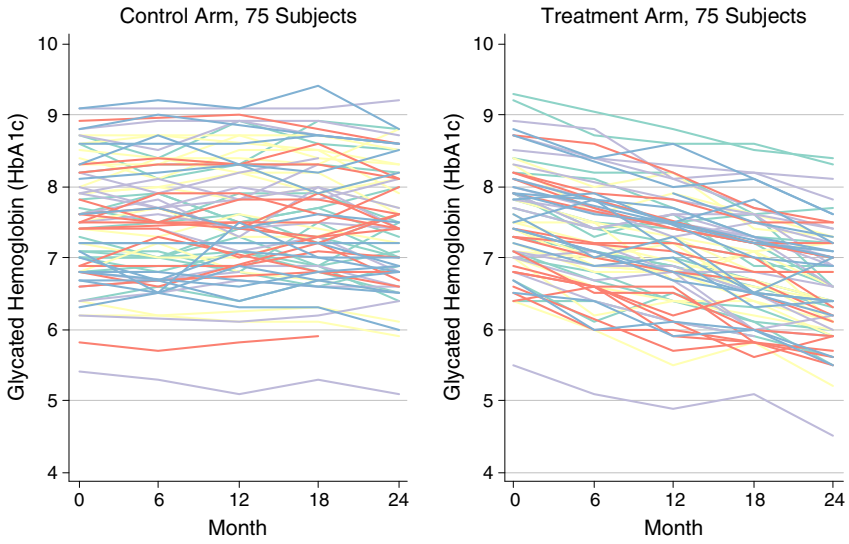


Figure 2. Multiple line (spaghetti) plot of glycated hemoglobin (data are simulated).

popular pasta dish if one has the right imagination. Spaghetti plots work well when there are relatively few subjects in the clinical trial, but they become difficult to interpret as the number of subjects increases.

We simulated data measuring glycated hemoglobin (HbA_{1c}) at baseline and 6, 12, 18, and 24 months for 150 subjects in a randomized controlled trial using a treatment and a placebo for the control arm. The average HbA_{1c} score was set to be 7.5% at baseline, and the treatment reduced HbA_{1c} by one percentage point over 24 months, on average.

We produced a spaghetti plot with the 75 control subjects in the left panel and the 75 treatment subjects in the right panel (Figure 2). We randomly assigned one of five colors – selected from the ColorBrewer [30] website – to the lines in an attempt to reduce the impact of overplotting. The plots show no trend in HbA_{1c} for the subjects in the control group, and a trend toward lower HbA_{1c} over the course of the trial for the subjects in the treatment group. However, there is so much overplotting that it is impossible to observe the individual level data except for a few patients with relatively extreme values.

Lasagna plots are an alternative approach to graph repeated measurement data for large numbers of subjects [31]. Lasagna plots depict the data for each individual subject as a row in a matrix plot. The data values are represented by gradations in color at each measurement interval. Let us look at one subject who was randomly assigned to the treatment group (Figure 3(a)). This subject had an HbA_{1c} measure of 7.1% at baseline. The score dropped to 6.7% at 6 months, and continued declining to 6.4%, 6.3%, and 5.9% in months 12, 18, and 24, respectively. The lasagna plot uses color to indicate these scores at each time point. In this case, we created categories that corresponded to the clinical interpretation of the HbA_{1c} test for diabetes risk, and we chose a set of colors from the selection of diverging color schemes available on the ColorBrewer website [30]. Other research has shown that even HbA_{1c} results in 5.5% to 6.0% range are associated with a statistically significant increase in the risk of diabetes. [32] We forfeit some of the precision of the underlying continuous variable by creating and plotting these categories. Researchers will need to use good judgment when considering the tradeoff between overplotting in a spaghetti plot and loss of information in a lasagna plot. Loss of information can be reduced by using more categories or a continuous scale for the color range.

In the full lasagna plot, each subject is depicted by an individual row, and the rows are layered to show all subjects. The rows can be arranged according to various criteria. In Figure 3(b), a case with 10 subjects, they are sorted first by whether the subject was in the treatment or control group and then by the subject's baseline HbA_{1c} score. Subjects in the control group were sorted in ascending order while those in the treatment group were sorted in descending order. This allows the reader to assess the symmetry of the two groups across the divider line.

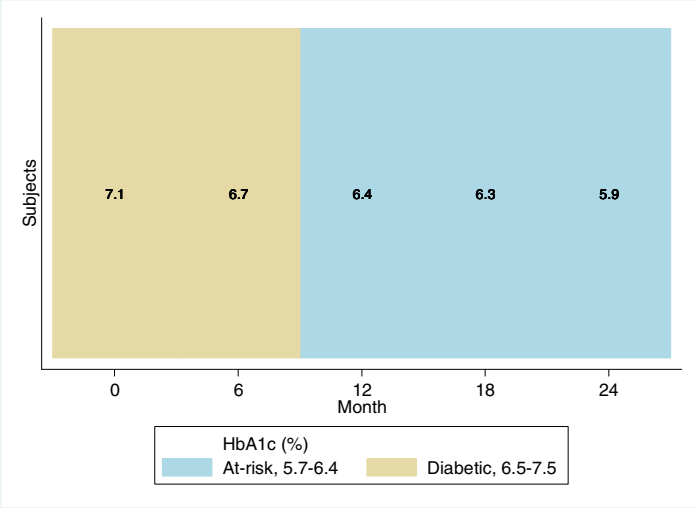
We believe the lasagna plot for all 150 subjects (Figure 3(c)) is much more informative than the corresponding spaghetti plot (Figure 2). Reading across each row allows the reader to explore the history for each individual subject. Most subjects in the treatment group trended toward lower HbA_{1c} scores as the trial progressed, but there was no clear trend in the control group. Reading down each column allows the reader to see the distributions of HbA_{1c} scores for the control and treatment groups at each measurement point. For example, the treatment and control groups had very similar distributions of HbA_{1c} scores at baseline, but in month 24, the distribution of HbA_{1c} scores for the treatment group had shifted to lower values. There was very little change in the distribution of HbA_{1c} scores for the control group.

We could not find any visual assessment studies comparing the lasagna plot with other ways of displaying repeated-measures data, so it is only our subjective opinion that it may be a useful option for some situations. We hope that this discussion stimulates new ideas about how to present repeated-measures data.

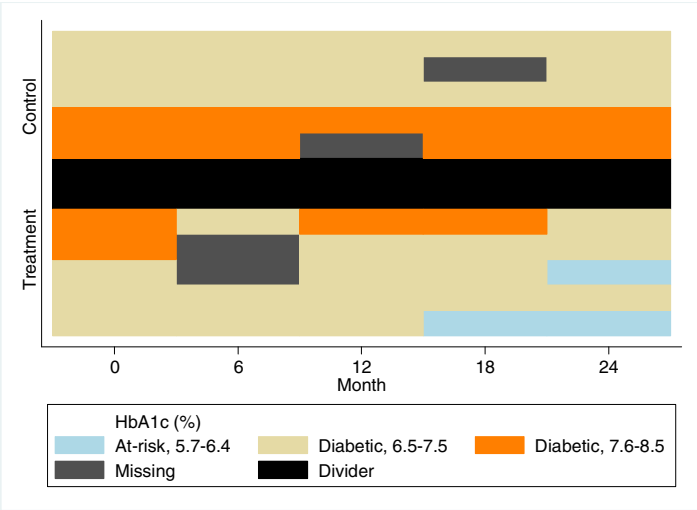
There are at least two limitations to lasagna plots. First, the lasagna plot requires readers to decode colors into numbers, which is not one of the best ways to encode numeric information in a statistical graph. Second, many lasagna plots use some binning of the underlying continuous data. In this case, we chose the cut points for the bins based on the ranges used for clinical interpretation of the HbA_{1c} results. Depending on the messages a particular dataset conveys, sometimes it will be better to make the bins as small as possible and use a color scheme that allows many bins. In other cases, it may be possible to represent the continuous scale of the underlying variable with an appropriate color scheme.

3.2. Bivariate plots for high density data

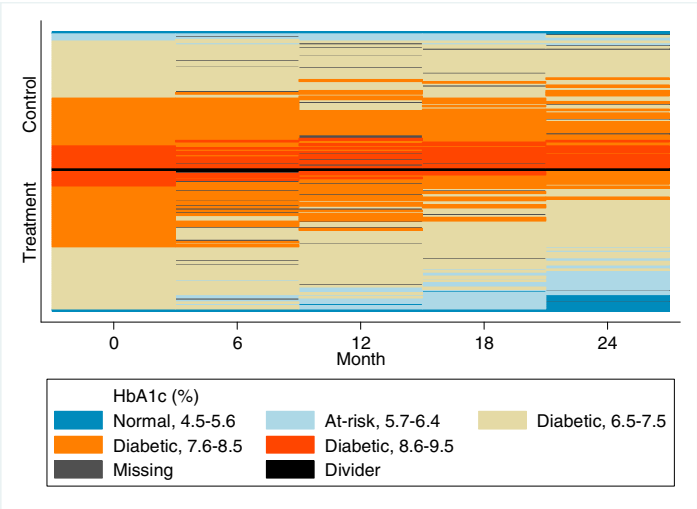
Scatterplots can also suffer from overplotting with large datasets. The plots in this section are based on the ADLBC dataset from the CDISC pilot and are available on the CTSPedia website [26, 33]. The



(a)



(b)



(c)

Figure 3. (a) Lasagna plot of glycated hemoglobin for a single subject (data are simulated). (b) Lasagna plot of glycated hemoglobin for 10 subjects (data are simulated). (c) Lasagna plot of glycated hemoglobin for 150 subjects (data are simulated).

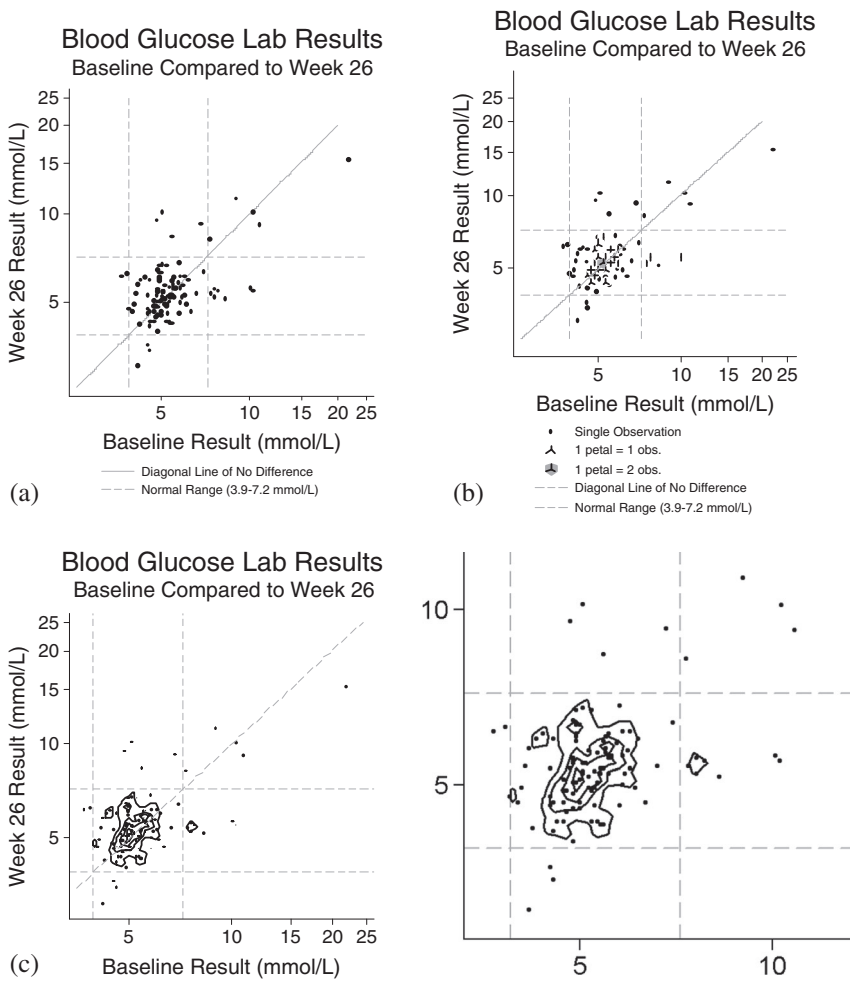


Figure 4. (a) Scatterplot of blood glucose levels at baseline and week 26. (b) Sunflower density plot of blood glucose levels at baseline and week 26. (c) Bivariate kernel density estimates of blood glucose levels at baseline and week 26 shown with contour lines. The left panel shows the full range of observed data values. The right panel shows an enlargement of the densest region of the data.

scatterplots examine the relationship between blood glucose levels (mmol/L) at baseline and week 26 in the clinical trial. We plotted the log of both variables to reduce the impact of a few extreme values, but we used the original units on the axes.

The scatterplot (Figure 4(a)) suggests a linear relationship between baseline blood glucose levels and the levels measured at week 26 in the trial. There is one extreme value at baseline of 22 mmol/L and week 26 of 15 mmol/L. While extreme, the value does not appear to be discrepant from the overall pattern in the data. There is a large cluster of observations in the general vicinity of baseline of 5 mmol/L and week 26 of 5 mmol/L, and the large number of observations in this area of the graph has produced some overplotting. A diagonal line representing no difference between baseline and week 26 and reference lines for the lower and upper bounds of the normal range of blood glucose levels have been added to aid in interpretation.

There are many techniques to reduce the impact of overplotting in bivariate graphs. Many of these techniques involve adjusting the markers for the data elements. For example, one may use smaller, hollow, or semitransparent markers to reduce the impact of overplotting. However, these methods may not be sufficient when there is a particularly dense area of data in a scatterplot.

We prepared two alternative examples for creating scatterplots with dense data. The first example, a sunflower density plot (Figure 4(b)), is an implementation of an approach that divides the graph region into hexbins, counts the number of observations that fall into each bin, and applies a visual scheme to display the count for each hexbin. The second example, as shown in Figure 4(c), uses contour lines to

show the bivariate kernel density. Heat maps, which use a color scale to encode the density of data points in regions of the graph, are another common approach to address this problem.

The sunflower density plot was proposed by Dupont and Plummer [34]. It divides the Cartesian plane into hexbins and uses flowers with petals to represent the number of data points in each bin. In this example (Figure 4(b)), a bin with more than one observation is shown with a petal that represents one observation. Bins with more than five observations are shown as a darker bin with each petal representing two observations. The sunflower density plot shows much more detail in the densest part of the plot than does the traditional scatterplot. This technique would be useful in situations where an understanding of the dense part of the plot is needed or with very large datasets.

For the second method, we estimated the bivariate density of the two variables and plotted the results using contour lines (Figure 4(c)) [35]. We included a scatterplot of the data to show the data outside of the densest regions. The contour lines show the greatest density in the normal range of HbA_{1c} at baseline and week 26, especially near the diagonal line of no difference. The left panel shows an enlargement of the region of the graph in which the data are the densest.

3.3. Network visualizations for large datasets

Network analysis investigates the complex relationships between the entities in social, biological, and other systems and supports the creation and analysis of network visualizations to represent them. The application of certain algorithms to the networks may reveal hidden patterns and emphasize specific dimensions of these relationships. Inspired by the principles of network theory and some applications to other areas, we propose the use of networks as a graphical method for the visual qualitative evaluation of medical outcomes in large datasets. Network analysis and visualization methods have been extensively described before [36, 37], so we do not discuss them here.

The Centers for Disease Control and Prevention report that 'Rotavirus is the leading cause of severe diarrhea and dehydration in young infants. Each year, rotavirus illness is responsible [for] an estimated 453,000 deaths among infants around the world.' The RotaShield® vaccine (RV), approved for use in the U.S. in 1998, was the first vaccine to prevent rotavirus [38, 39]. However, RV was reported to be associated with intussusception (IS) in 1999 [40]. IS is the invagination of bowel into the lumen of adjacent bowel and is the most common cause of intestinal obstruction in infants. It is frequently preceded by a viral infection in children and presents with obstructive and other gastrointestinal (GI) symptoms [41]. Recent studies of second-generation rotavirus vaccines show a much smaller association with IS [42].

We conducted an exploratory analysis to determine whether network visualizations could also discover this known safety signal and provide a richer picture of the IS pattern. We retrieved all the safety reports submitted to the U.S. Vaccine Adverse Event Reporting System for RV from its licensure in August 1998 to its withdrawal in October 1999. The elements of the reports (i.e., the vaccines names and AE codes) and their co-occurrence are represented as the nodes and edges (or connections) in the network shown in Figure 5(a). The edges are weighted according to the number of reports containing the corresponding pair of elements. Network visualizations can be organized in different ways to study various features of the data. We created several network analysis visualizations to explore patterns in the data.

We constructed the network in Figure 5 using the force-directed layout algorithm. The basic principle of this algorithm is that it places the most connected nodes in the center of the network. RV and other vaccine nodes as well as IS, GI, and other AE nodes are among them.

However, the network in Figure 5(a) is a relatively large structure and does not allow the complete evaluation of the most connected vaccine and AE central nodes. We used the islands algorithm and reduced the network by identifying the top 30 nodes in terms of their 'connectedness' [43, 44]. The number of nodes for the reduced network was determined after a subjective examination of the current visualization; the selection of fewer nodes might have included part of the pattern whereas more nodes would simply add more noise to it. The resulting 30-node reduced subnetwork supports a more focused evaluation of the highly connected nodes through the creation of various network layouts that appear in panels b, c, d, and e. Each layout emphasizes different aspects of the 30-node subnetwork and supports the clinical interpretation of the visual findings.

We initially applied the self-organizing map (SOM) layout algorithm that places nodes such that they appear near to and far from nodes with which they are highly connected and not connected at all, respectively. The visualization in Figure 5(b) emphasizes the co-administration of RV with other vaccines (overlapping square nodes in the center of the network image) and suggests the existence of an IS pattern

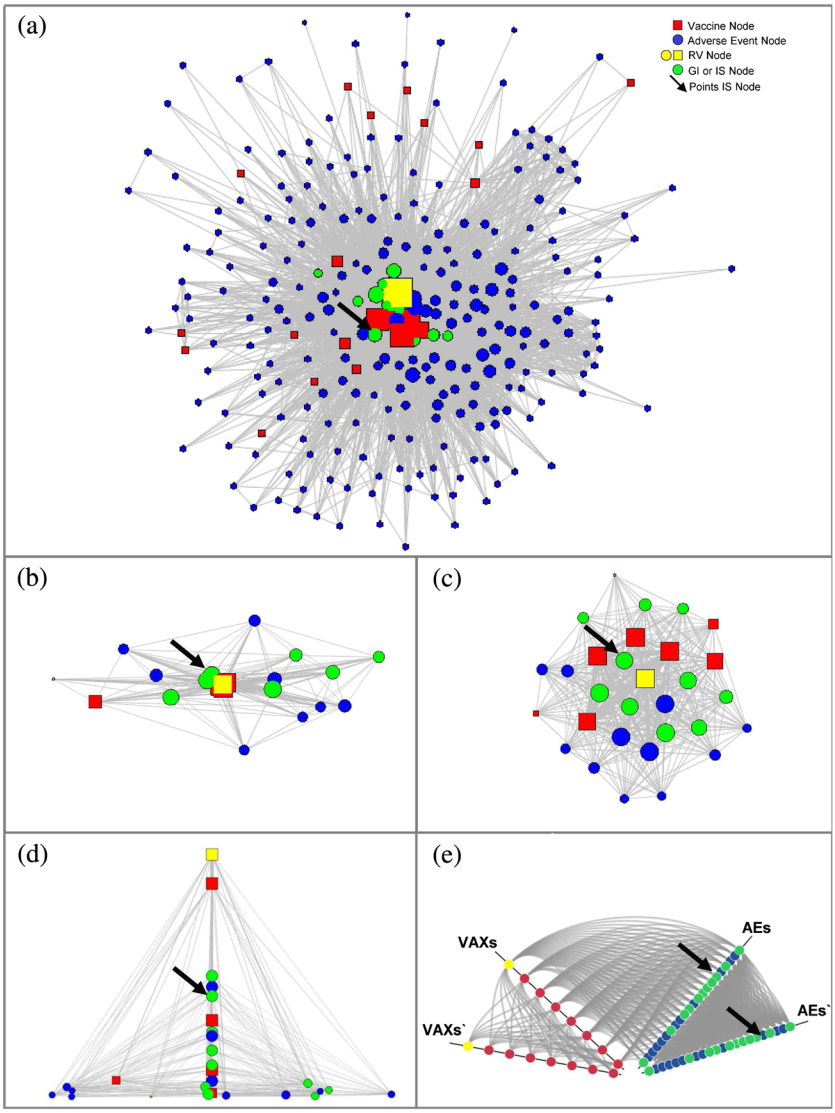


Figure 5. Network visualizations for the administration of RotaShield® vaccine (RV) and its tight association with intussusception (IS) and other gastrointestinal (GI) adverse events. We start with the force-directed layout for the full network, shown in (a), to evaluate whether there is a safety pattern in the data deserving further exploration. The dense region in the center of the network appears to contain a constellation of tightly connected nodes, potentially representing a pattern. This is not quite clear in the large visualization though. It is therefore necessary to isolate the most connected nodes in a reduced network that will help us verify the existence and the nature of the pattern and, particularly, the contributing adverse events (AEs) and the vaccine product(s) associated with it; the four layouts in (b), (c), (d), and (e) highlight these important aspects. The self-organizing map layout in (b) tries to place nodes so that distances are meaningful and informative of the node connectivity – closest should be interpreted as more connected. The self-organizing map shows at least five overlapping GI AE nodes (the IS node is among them) associated with a number of overlapping vaccine nodes. This indicates the existence of a potential pattern; however, the pattern and the associated vaccine product(s) are hidden in the constellation of visual elements. The force-directed layout in (c) verifies the existence of a GI-related pattern and its tight connection with a single vaccine node (that is the RV node). The islands layout in (d) is more specific because it places both the RV and the IS nodes on the top of the network and, thus, verifies the occurrence of IS following the administration of RV vaccine. Finally, the hive plot layout in (e) places AEs and vaccines on two separate axes and arranges them based on the number of other nodes they are connected to – the more connections they have the farther out they are on the axis. The lines between vaccines and AEs axes show the connections between vaccines and AEs. Their clone mirror axes (vaccines and AEs, respectively) illustrate the connections between the entities of the same type (vaccine-to-vaccine and AE-to-AE). VAX, vaccine.

(IS and other GI nodes aggregate in a dense group). The IS node, mainly, and one other GI node, secondarily, are tightly connected to the overlapping vaccine nodes. However, the actual vaccine node that is directly associated with IS is hidden in the SOM layout.

We therefore re-ran the force-directed algorithm on the 30-node reduced subnetwork. This resulted in the visualization shown in Figure 5(c) that verifies the existence of IS pattern (also found in the SOM layout) and reveals the tight connection of IS to the RV vaccine node. RV is also connected to the other GI nodes of the IS pattern and appears at the center of the topology.

To further evaluate the role of RV, we created the islands layout (Figure 5(d)) that organizes nodes vertically and horizontally based on their interconnectivity. The RV node sits on the top of the layout with one of its co-administered vaccines, and the IS-related nodes are on the same vertical axis, tightly connected to RV. The AE node (node in blue) between the IS and the 'Diarrhea' node is the commonly reported 'Pyrexia' that appears in the majority of Vaccine Adverse Event Reporting System reports and does not affect our clinical interpretation of this layout.

The previous visualizations focused on the multiple interactions between all the entities of the dataset irrespective of their type (vaccine or AE). The hive plot layout (Figure 5(e)) groups the nodes by type and arranges them along radially distributed linear axes based on a selected node property, such as the number of other nodes that node is connected to (in network theory terms, this is defined as 'degree centrality'). Hive plots also allow the creation of clone axes to illustrate the connections between the nodes of the same type. The AE axis and its clone in Figure 5(e) show the strong connections between IS and the other GI AEs. They are all connected to the RV node, which is the top vaccine node in terms of degree centrality. The AE part of the hive plot shows that the AE axes are mainly occupied by the IS and the other GI nodes that also co-occur in many reports – this is indicated by the multiple lines linking the nodes on the AE to the nodes on the AE axis (or vice versa). The vaccine part of the plot has fewer nodes, because we have fewer vaccines than AE nodes in the network and even fewer co-administered vaccines. The lines that bridge the two parts show that the top vaccine nodes are connected with more AE nodes than the vaccine nodes that are closer to the origin of the axes.

The hive plot was created by the jhive tool (developed by Martin Krzywinski; available at http://www.hiveplot.net/) and the remaining four by the U.S. FDA Adverse event Network Analyzer [45].

4. Improving graph quality by applying best practices

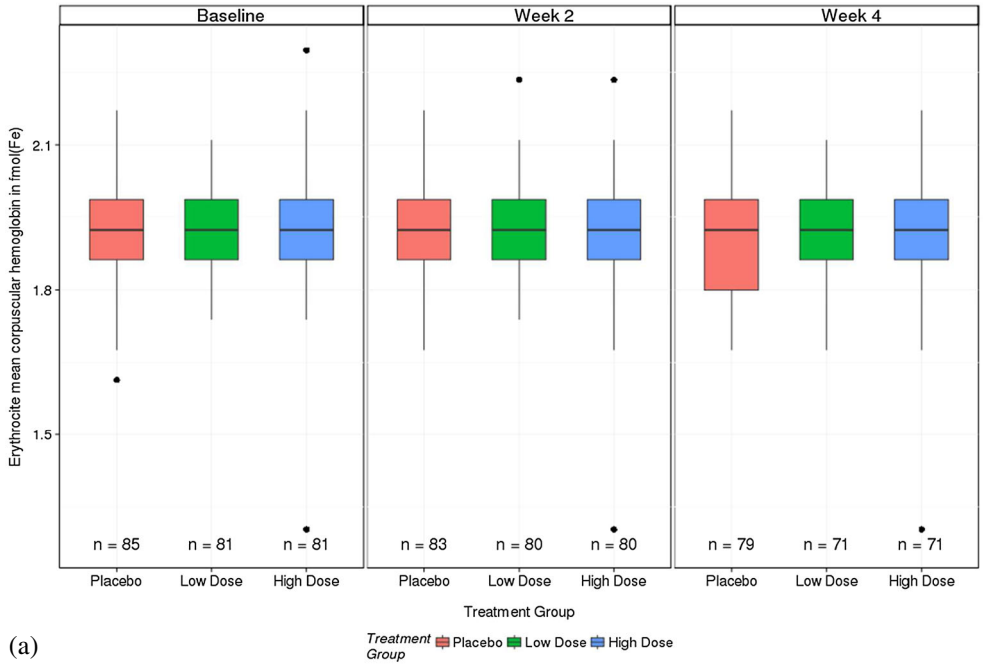
The General Principles subteam of the Safety Graphics Working Group recommends nine best practices for generating statistical graphics in clinical trials (see Appendix). Considering the compelling nature of a visual display to its audience, we encourage statisticians to adhere to these principles just as carefully as they consider the assumptions of statistical models, multiplicity adjustments, and other statistical principles. In addition, statistical graphs, like any other statistical method, often rely on certain assumptions. For example, graphs that only show change from a baseline value assume that the baseline values contain no relevant information. Graph designers should clearly communicate any assumptions used and carefully consider whether the assumptions are valid. The examples in the succeeding text demonstrate many of these principles.

The Safety Graphics Working Group used these principles *as a checklist* to review each of the graphs recommended on the wiki (see Appendix). We recommend a similar approach to team members and/or statisticians when they design and review graphs. Often, adjustment may be needed to improve clarity after reviewing graphs that display the actual data.

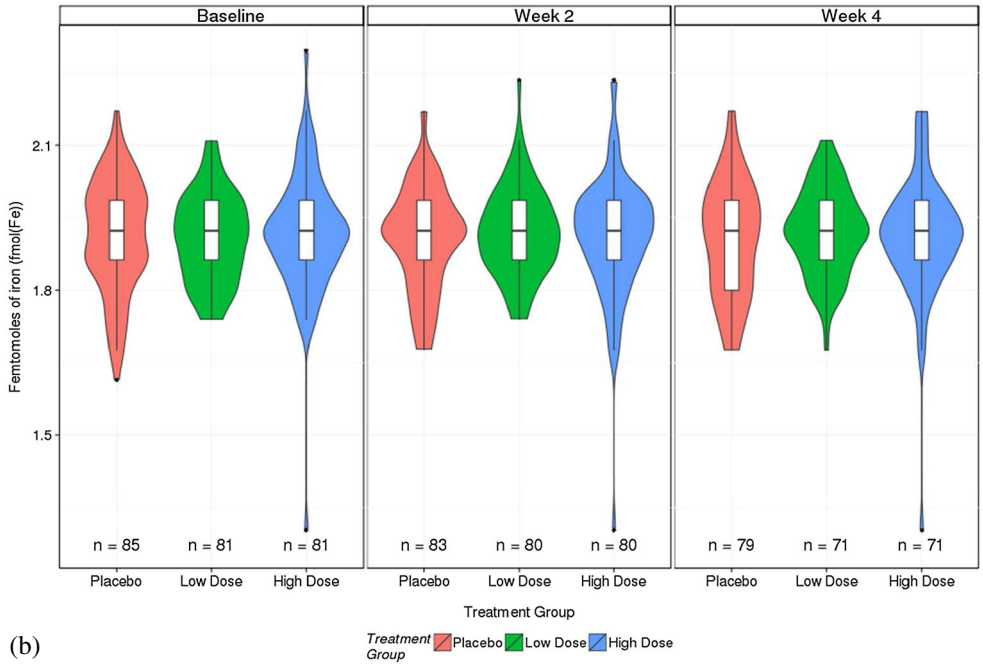
In this section, we demonstrate good graphics design principles through several examples, which illustrate the following principles:

(1) tailor each graph to its primary communication purpose,
(2) maximize data-to-ink ratio,
(3) bring items the reader needs to compare close together,
(4) use the simplest plot that is appropriate for the information to be displayed, and
(5) avoid misleading visual perception.

For the examples, we begin most with a suboptimal graph, followed by a graph that enhances the reader's ability to visually decode its key message(s). Good graphs, like good writing, require a few rounds of revisions. We demonstrate in the examples the type of thinking needed to make almost any graph a good one.



(a)



(b)

Figure 6. (a) Boxplot of erythrocite mean corpuscular hemoglobin in fmol(Fe) over time by three treatment groups The extent of the box is the interquartile range (IQR), with the median inside. Whiskers extend to min and max if the latter are within 1.5 IQR from the box. Otherwise, whiskers just extend to 1.5 IQR from the box, and any values beyond 1.5 IQR from the box are displayed as outliers (single dots). The data are a subset of the ADLBH dataset, which can be found at http://phuse-scripts.googlecode.com/svn/trunk/scriptathon2014/data/. (b) Violin plots of the same data as used in (a).

4.1. Tailor each graph to its primary communication purpose

Boxplots are often used to display quantitative safety data such as laboratory values and vital signs. Figure 6(a) shows a boxplot of a lab parameter – erythrocyte mean corpuscular hemoglobin, fmol(Fe), over time and by treatment group.

The violin plot adds a kernel density plot to each side of the boxplot and thus more details of the data distribution. A violin plot of the same data (Figure 6(b)) shows that the shapes of the distributions look somewhat different over time and by treatment group.

A boxplot shows more distributional detail than a dot plot. The higher level of detail in a violin plot allows the graph reader to assess distributional shapes not captured by simple summary statistics afforded by the boxplot. As illustrated in Figure 1(a), one of the important factors that influences which type of graph to use is this level of distributional detail. This will vary depending on the level of detail needed to convey the relevant clinical messages about the variable(s) of interest.

4.2. Maximize data-to-ink ratio and bring items the reader needs to compare closer together

Imagine an early clinical trial that wishes to assess the effects of multiple doses of an experimental treatment (drug low, drug mid, and drug high) on systolic blood pressure. To assess the effects of the treatment, the mean at the end of treatment was calculated along with its associated 95% CI for each of the doses as well as the control. Also of interest is whether the effect is similar for males and females.

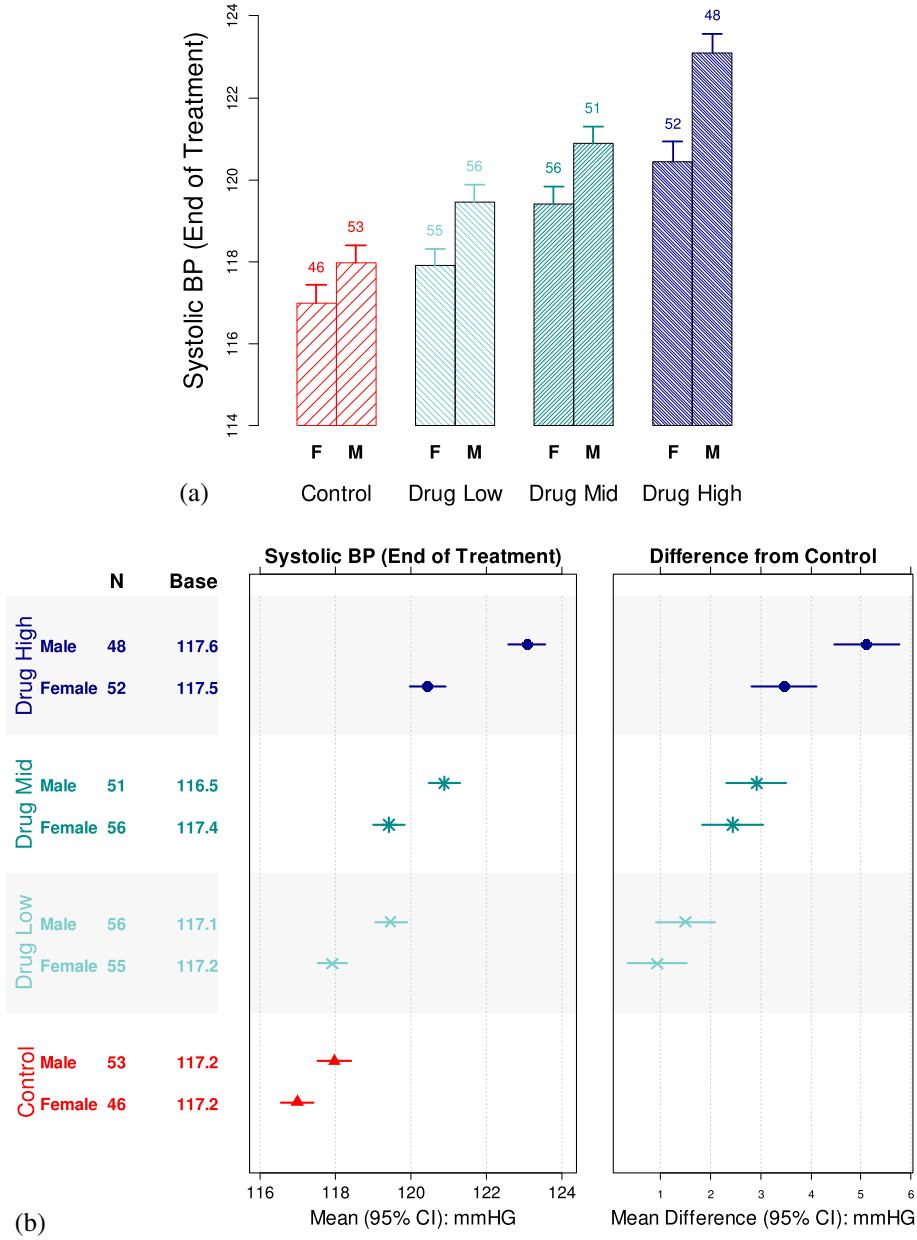


Figure 7. (a) Barplot of systolic blood pressure at end of treatment (not recommended). Data are hypothetical. (b) Multipanel dot plot systolic blood pressure (recommended). BP, blood pressure; F, female; M, male.

Figure 7(a) depicts a common way such data are graphed by the use of a barplot with an error bar at the top of the bar as well as the sample size above the error bar. Presenting the error bar and text above the error bar creates an optical illusion by creating a visual 'addition' to the actual data value itself (the height of the bar). Using a barplot with error bars fails to provide the statistical parameter depicted and does not indicate what the tail on top of each bar represents (these are the mean and the upper bound of the 95% CI). Note that the bar height represents the mean, and there is no additional information inherent in the bar itself other than this single numerical value; therefore, the data-to-ink ratio is not maximized. Barplots also fail to clearly distinguish the lower bound of the CI as this is masked by the bars in the figure.

Such a graph design impedes on the ready interpretation of the data. As is the case in many clinical trials, the quantity of interest is the comparative differences between the experimental treatment and that of the control. This requires one to move the eye from right to left to assess the difference between the experimental dose and control, without conveying the statistical uncertainty for this comparison – a task made more difficult by also assessing effects of gender.

A multipanel dot plot was used to improve the graphic design. Figure 7(b) depicts both the mean systolic blood pressure at the end of treatment and the associated 95% CI in the first panel along with the mean difference from control plus the associated 95% CI for each of the experimental treatment arms. The dot plot uses a single plotting character to depict the estimate of the parameter of interest with lines corresponding to the 95% CI, a much more effective data-to-ink ratio (resulting in graph clarity and unobstructed data patterns). The difference between the experimental dose and that of the control is easily decoded in the right-hand panel of the dot plot as it is clearly indicated which experimental doses (and gender) have lower bounds greater than 0 (values greater than 0 imply increases in systolic blood pressure). While not necessarily the case with this example, an additional advantage of using a dot plot over a barplot is that a dot plot allows for the presentation of a greater number of categories.

4.3. Use graphs that augment traditional methods of displaying data

A typical presentation of adverse events of special interest (AESIs), which are generally limited to a finite set of events, summarizes event and/or incidence rates in tabular form for each of the treatment arms.

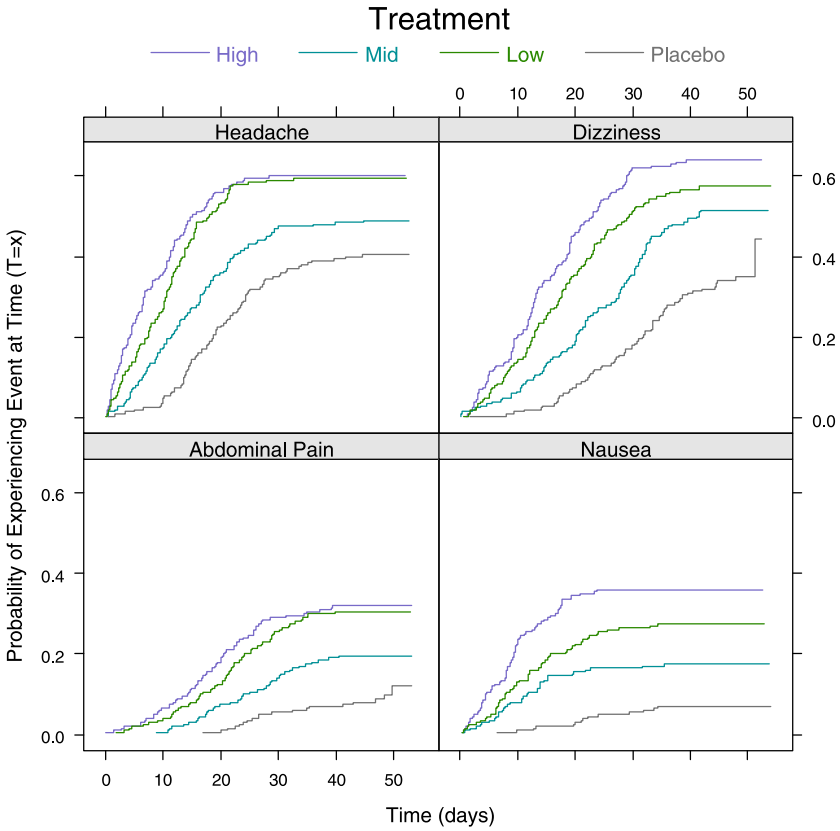


Figure 8. Time-to-event visualization of common adverse events (recommended).

An alternate, more statistically driven approach would be to model the event rates using time-to-event methodologies. The modeling of these AESIs can be presented using multipanel figures with each AESI serving as a panel in the graph with grouping of the treatments distinguished by either color or line type.

For example, suppose we have four AESIs: headache, dizziness, abdominal pain, and nausea. Three doses are explored along with a placebo control. Each of the AESIs is modeled using time-to-event methods with a Kaplan–Meier curve constructed for each AESI. Figure 8 depicts the Kaplan–Meier plot with a panel for each AESI. With such a figure, one is able to identify differences in the frequency of the event between treatment groups, which would be captured in a tabular reporting of the event rates. However, the figure aids in the understanding when the AESI might occur at different rates or onset times between the treatment groups.

It is important to take into account different safety outcomes in the evaluation of a drug. Creating dashboard-like figures such as this allows the reader to detect the absence/presence (and understand the magnitude) of effect (including if dose-dependent or not).

4.4. Avoid misleading visual perception

The human brain is better at processing some patterns than others. Statisticians Cleveland and McGill [14] conducted experiments to understand human graphical perception capabilities and their application to the development of graphical methods. They found that volume and color rank low on the list of what humans perceive well.

Pie charts are ubiquitous in clinical research, which is likely a reflection of their use in the general press. While they may be easy to create, the human brain is relatively poor at quantifying angles and areas and some colors. As an example, which is the largest area in Figure 9(a), A, C, or F? Which is the smallest? When a group of 38 statisticians and programmers were asked to rank the letters from least to greatest, only 5% were correct.

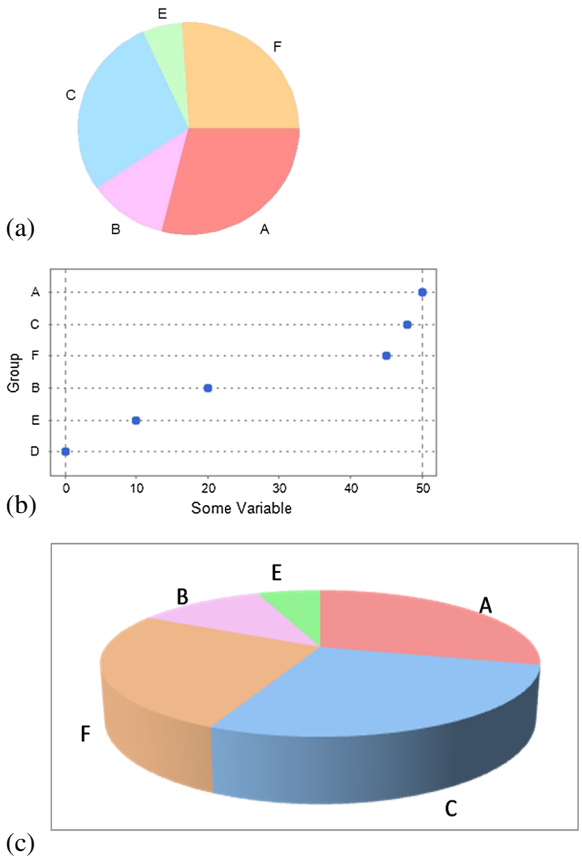


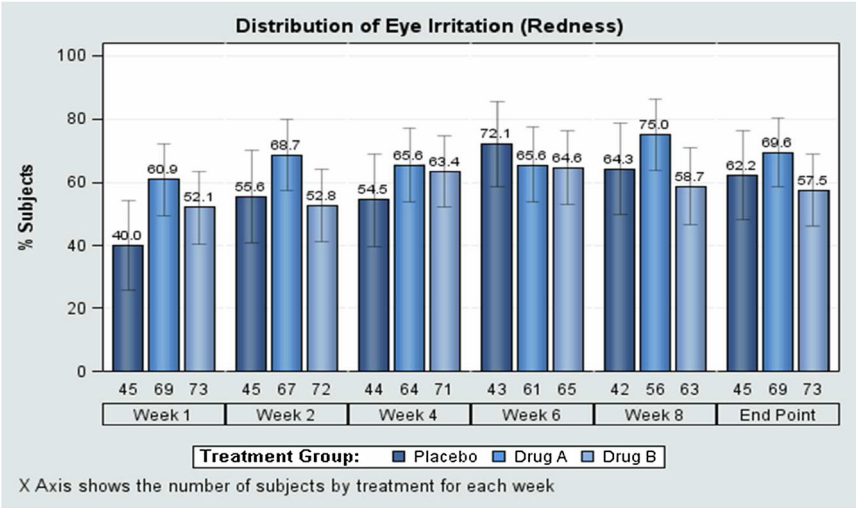
Figure 9. (a) A pie chart used in an informal study to assess how well data can be read from it (not recommended). (b) Dot plot using the same data as (a) (recommended). (c) Pseudo three-dimensional pie chart (using the same data as (a)) (not recommended).

The dot plot in Figure 9(b) displays the same data as in Figure 9(a). Did you miss letter D? 92% of those who ranked the letters in Figure 9(a) did. If it is important that the value for D was 0, then most would have missed it from the pie chart.

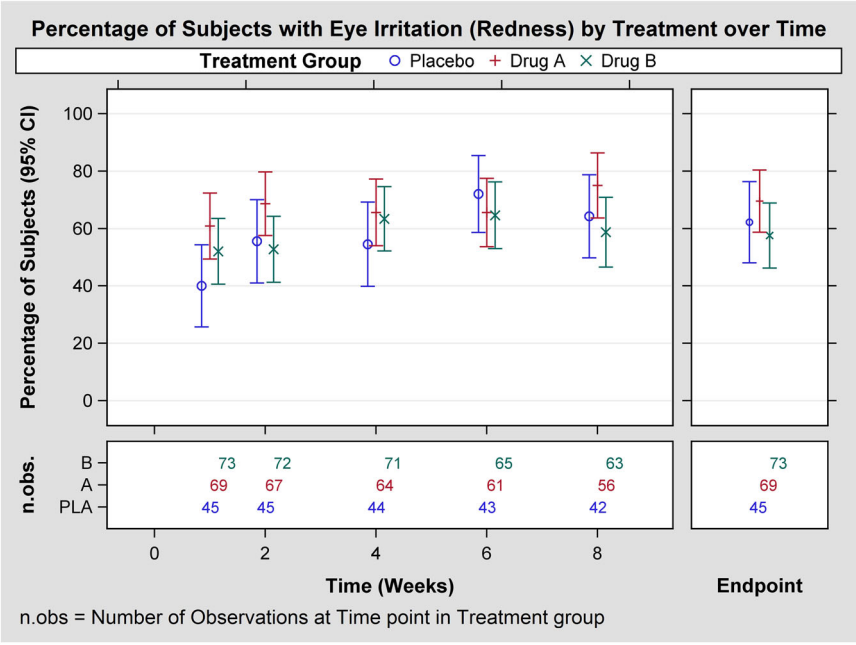
If reading pie charts is a challenge, there is something even worse, the pseudo three-dimensional pie chart in Figure 9(c), which provides the distorted view that C is larger than A.

4.5. Appropriately handle quantitative x-axis data and avoid misinterpretation by separating endpoint data from 'over-time' data

This example further illustrates the good graphing principle 'maximize the data-to-ink ratio'. Figure 10(a) shows the percentage of subjects with eye redness over time in a study for three treatment groups. The figure is pleasing to the eye but has several possible areas for improvement. Much of the ink used in Figure 10(a) does not aid the reader in their interpretation of the data. In fact, the important message of the



(a)



(b)

Figure 10. (a) Bar chart of distribution of eye irritation. (not recommended). The data are given in http://support.sas.com/kb/39/166.html along with the code we used to create this figure. The data are percent of subjects with eye irritation at five time points – weeks 1, 2, 4, 6, and 8 and at endpoint. (b) Dot plot of distribution of eye irritation (recommended) using the same data as (a).

plot is somewhat obscured. The main information is the percent of subjects, with CIs (or some measure of variability – it is not clear what the measure of variability is). However, all the ink in the bars obscures this information. Only the height of the bars is important, rather than the filled bar parts themselves.

Another problem with this graph is the *x*-axis represents the continuous variable of time as a categorical variable. The eye irritation was measured at weeks 1, 2, 4, 6, and 8, but the spacing on the *x*-axis makes it appear that they were measured at equal time intervals. Additionally, the 'End Point' is not clearly distinguished from the data at the specific weeks.

A minor issue is the choice of colors. The plot indicates that the three treatment groups are placebo, drug A, and drug B. Figure 10(a) uses sequential colors (light to dark), which might be a better choice for treatment groups that progress from low to high such as placebo, low dose, and high dose. In this case, it would be better to choose colors that suggest a qualitative difference between the groups. Qualitative schemes do not imply magnitude differences between legend classes, and hues are used to create the primary visual differences between classes. (We do recognize, however, that depending on what drug A and drug B are, a sequential scheme might be appropriate.) Graphs that use an appropriate color scheme will communicate their messages more effectively. There are many resources to help statisticians choose effective color schemes, such as ColorBrewer [30].

Figure 10(b) has several improvements over Figure 10(a). First, by removing all the excess ink and only showing the point estimates of percent of subjects with CIs, it is much easier to see the trends. Next, time in weeks is a quantitative variable, rather than categorical. Now, weeks 1 and 2 are visually closer than weeks 2, 4, 6, and 8. Furthermore, endpoint is clearly delineated from the time in weeks. Figure 10(b) also indicates that the measure of variability is 95% CI, fulfilling the good graphics principle of having all information needed for interpretation on the plot itself. A qualitative color scheme was chosen using ColorBrewer [30].

5. Conclusions

Through new flow diagrams for selecting the optimal graph type (Figure 1(a–c)), description/encouragement of some complex graphs, and with selected illustrations and application of best practice, this manuscript recommends more effective graphics to convey key data messages. While these examples do not account for all possibilities, the flow diagram provides considerations that the conscientious statistician can use when designing a new graph. The graphics principles themselves serve as a convenient checklist when reviewing the initial draft.

Effective statistical graphs require careful consideration of several important principles that may sometimes be in conflict. However, with careful consideration of best practices and with feedback from the other disciplines on your team, a well-constructed graph will likely be the result. It may take additional time, but in the end, effective statistical graphics are fast and powerful in their ability to communicate key findings to the decision-makers who base their decisions on the analysis in medical research reports, regulatory applications, and publications.

Appendix: Nine Graphics Principles [46]

(1) **Content** Every graph should stand on its own.

   (1) It should tell its story without a need for detailed explanatory text or supporting documents. Consider annotations and self-explanatory symbols rather than a remote legend.
   (2) It should be clear, effective, and informative for the intended audience.

(2) **Communication** Tailor each graph to its primary communication purpose.

   (1) What insight is the graph intended to convey? Is it intuitive?
   (2) Avoid packing too much information into a single display and distracting from the main message.

(3) **Information** Maximize the data-to-ink ratio.

   (1) Each spot of ink should be necessary for imparting the main message.
   (2) Do not clutter a graph with what you don't need. Less is more.
   (3) Avoid overplotting.

(4) **Annotation** Provide legible text and information.

   (1) Position annotation (including legends) so that it aids interpretation and does not distract from the message.
   (2) Use legible font that can be read without eye strain or a great deal of effort. Consider the format (presentation or document).

(5) **Axes** Design axes to aid interpretation of a graph.

   (1) Scale axes to show the interesting features of the data; for example, for longitudinal data, use time (on a continuous scale) instead of visit number (on an ordinal scale).
   (2) Give careful consideration to inclusion of the zero of each axis; if excluded, ensure its absence is clearly sign-posted.
   (3) Avoid crowded axes.
   (4) Use the same axis scales on graphs that need to be compared.
   (5) Choose the appropriate style of axes. For example, select between a box, X and Y axes, X only, Y only; consider grid lines; ensure intelligent placing of tick marks.
   (6) If the nature of the data suggests the shape of the graphics, follow that suggestion; otherwise, use horizontal graphics about 50% wider than tall.

(6) **Styles** Make symbols and plot lines distinct and readable.

   (1) Choose plot symbols with simple, familiar shapes and intuitive interpretation (e.g., 'A' for active and 'P' for placebo) which eliminates the need for a remote legend.
   (2) If a graph is to be displayed by projection onto a screen, or in a poster, use thick lines, large symbols, and large fonts to achieve legible display.
   (3) Where possible and appropriate, data representations (such as styles of symbols, lines, and bars) should have the same meaning across all similar graphs within a package; for example, if one line graph uses a solid blue line to represent placebo, all graphs in the package should use a solid blue line for placebo.

(7) **Colors** Make use of color if appropriate for the medium of communication.

   (1) Use color only when it decodes information. When color is used, choose contrasting and clearly visible colors; avoid yellow, and contrasts with red, green, or brown, which are difficult for people with color-deficient vision.
   (2) If a graph may be viewed in black and white, ensure that all distinctions made by color are also made by other features such as symbols and line-styles.
   (3) For black-and-white media, make use of line-styles (dashing and gray levels) that are easy to distinguish.
   (4) Design backgrounds to set off the graph, not compete with it.
   (5) Choose area fills that are distinct but compatible.
   (6) Make secondary plot lines lighter in weight, color, or style.
   (7) Keep reference lines and grids distinct from other data lines.
   (8) ColorBrewer is an excellent reference for choice of colors.

(8) **Techniques** Use established techniques to clarify the message.

   (1) Show causality: when a causal relationship exists between variables, make sure it is easily discernible from the graph.
   (2) Make comparisons from a common reference point, carefully considering whether the assumptions are valid and clearly communicating those assumptions.
   (3) Sort categories according to relevant features of the data.
   (4) Do not introduce spurious dimensions to a graph, as they reduce clarity.
   (5) Combine multiple images into a single display when information needs to be presented together.
   (6) When a graph summarizes data at an aggregate level, always plot estimates of variability in the data.

(9) **Types of plots** Use the simplest plot that is appropriate for the information to be displayed (see Select the Right Graph for My Question).

(1) To show a distribution of values, use whichever form is most appropriate: rug plot, strip plot, dot plot, boxplot, histogram, CDF plot, or more specialized display.
(2) Use scatterplots and line plots to show association between a pair of variables, thinking carefully about the representation of variability of actual data.
(3) Use trellis displays to show changes in association between a pair of variables with respect to a third variable.

Adapted from: GlaxoSmithKline Graphics Principles (used with permission).

Acknowledgements

The authors wish to thank Peter Lane (Australian National University) and Mark Jones (Amadeus Software) for their influence and efforts while employed at GlaxoSmithKline on their graphics initiative (and in particular, the graphics principles and pie chart example, respectively). The GSK graphics initiative was also a motivator for the creation of the FDA/Industry/Academia Safety Graphics Wiki team. We wish to thank the CTSA Consortium, specifically the Biostatistics/Epidemiology/ Research Design Key Function Committee (BERD), for their website support of the Safety Graphics initiative. We also wish to thank our fellow members of the Safety Graphics Wiki team, particularly the other members of the Graphics Principles subteam, Frank Harrell (Vanderbilt University), Qi Jiang (Amgen), Richard Anziano (Pfizer), Mary Banach (Vanderbilt University), Larry Gould (Merck), and Andreas Krause (Actelion). We thank Rebeka Tabby (Lilly) for creating the box and violin plots. Finally, thanks to the reviewers from each of our organizations and especially the reviewers of this journal for their thoughtful and helpful comments.

Disclaimer: *The opinions and information in this review article are those of the authors, and do not represent the views and/or policies of the U.S. Food and Drug Administration.*

References

1. Krzywinski M, Cairo A. Points of view: storytelling. *Nature Methods* 2013; **10**:687.
2. Robbins N. *Creating More Effective Graphs* (2nd edn). Chart House: Wayne, NJ, 2013.
3. Kirk A. *Data Visualization: A Successful Design Process*. Packt Publishing: Birmingham, UK, 2012.
4. Cairo A. *The Functional Art: An Introduction to Information Graphics and Visualization (Voices That Matter)*. New Riders: Berkeley, CA, 2013.
5. Bertin J. *Semiology of Graphics: Diagrams, Networks, Maps*. University of Wisconsin Press: Madison, WI, 1983.
6. Friendly M. *Visualizing Categorical Data*. SAS Institute Inc: Cary, NC, 2000.
7. Card S, Mackinlay JD, Shneiderman B. *Readings in Information Visualization: Using Vision to Think*. Academic Press: San Diego, CA, 1999.
8. Cook D, Weisberg S. *Applied Regression Including Computing and Graphics*. John Wiley & Sons: New York, NY, 1999.
9. Wickham H. *ggplot2: Elegant Graphics for Data Analysis (Use R!)*. Springer: New York, 2009.
10. Tufte ER. *The Visual Display of Quantitative Information* (2nd edn). Graphics Press: Chesapeake, CT, 2001.
11. Cleveland WS. *The Elements of Graphing Data* (2nd edn). Hobart Press: Summit, NJ, 1994.
12. Cleveland WS. *The Collected Works of John Tukey: Graphics 1965–1985*, Vol. V. Chapman & Hall: New York, NY, 1988.
13. Tukey J. Data-based graphics: visual display in the decades to come. *Statistical Science* 1990; **5**:327–339.
14. Cleveland WS, McGill R. Graphical perception: theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association* 1984; **79**:531–554.
15. Tukey J. *Exploratory Data Analysis*. Addison-Wesley: Reading, MA, 1977.
16. Singer PA, Feinstein AR. Graphical display of categorical data. *Journal of Clinical Epidemiology* 1993; **46**(3):231–236.
17. Bradstreet TE. *Graphical Excellence - The Importance of Sound Principles and Practices for Effective Communication* 52nd Session. International Statistical Institute: Helsinki, 1999.
18. Pocock SJ, Travison TG, Wruck LM. Figures in clinical trial reports: current practice & scope for improvement. *Trials* 2007; **8**:36–53.
19. Cleveland WS. A model for studying display methods of statistical graphics. *Journal of Computational and Graphical Statistics* 1993; **2**(4):323–343.
20. Gelman A, Unwin A. Infovis and statistical graphics: different goals, different looks. *Journal of Computational and Graphical Statistics* 2013; **22**(1):2–28.
21. Gelman A, Pasarica C, Dodhia R. Let's practice what we preach: turning tables into graphs. *The American Statistician* 2002; **56**(2):121–130.
22. Connor JT. Statistical graphics in AJG: save the ink for the information. *American Journal of Gastroenterology* 2009; **104**:1624–1630.
23. Jacoby WG. *Statistical Graphics for Univariate and Bivariate Data (Quantitative Applications in the Social Sciences)*. Sage Publications: Thousand Oaks, CA, 1997.
24. Puhan M, ter Reit G, Eichler K, Steurer J, Bachman LM. More medical journals should inform their contributors about three key principles of graph construction. *Journal of Clinical Epidemiology* 1017; **59**(10):2006.
25. Krause A, O'Connell M. *A Picture is Worth a Thousand Tables: Graphics in Life Sciences*. Springer Science: New York, 2012.

26. The Clinical Trials Safety Graphics Home Page. Available at: https://www.ctspedia.org/CTSpedia/StatGraphHome [accessed 02 September 2013].
27. Data visualization flow diagram from the Clinical Trials Safety Graphics Home Page. Available at: https://www.ctspedia.org/CTSpedia/SelectRightGraph [accessed 03 September 2013].
28. Amit O, Heiberger R, Lane P. Graphical approaches to the analysis of safety data from clinical trials. *Pharmaceutical Statistics* 2008; **7**:20–35.
29. Data visualization flow diagram. Available at: https://www.ctspedia.org/CTSpedia/SelectRightGraph [Accessed on 1 September 2013].
30. Brewer C, Harrower M. Color Brewer: color advice for maps. Available at: http://colorbrewer2.org/ [accessed 22 September 2013].
31. Swihart BJ, Caffo B, James BD, Strand M, Schwartz BS, Punjabi NM. Lasagna plots: a saucy alternative to spaghetti plots. *Epidemiology* 2010; **21**:621–625.
32. Selvin E, Steffes MW, Zhu H, Matsushita K, Wagenknecht L, Pankow J, Coresh J, Brancati FL. Glycated hemoglobin, diabetes, and cardiovascular risk in nondiabetic adults. *New England Journal of Medicine* 2010; **362**:800–811.
33. Dataset from CDISC SDTM/ADaM Pilot Project. Available at: https://www.ctspedia.org/CTSpedia/DataSetPageCDISC [accessed 03 September 2013].
34. Dupont WD, Plummer WD. Density distribution sunflower plots. *Journal of Statistical Software* 2003; **8**:1–5.
35. Gallup JL, Baum CF. *BIDENSITY: Stata Module to Produce and Graph Bivariate Density Estimates*, Statistical Software Components S457576. Boston College Department of Economics: Boston, MA, 2013.
36. DiBattista G, Eades P, Tamassia R, Ioannis TG. *Graph Drawing: Algorithms for the Visulaization of Graphs*. Prentice Hall: Upper Saddle River, NJ, 1999.
37. Newman MEJ. *Networks: An Introduction*. Oxford University Press: Oxford, UK, 2010.
38. Centers for Disease Control and Prevention. *Vaccine Safety: Rotavirus*, 2014. Available at: http://www.cdc.gov/vaccinesafety/Vaccines/RotaVSB.html [Accessed on 11 June 2015].
39. Centers for Disease Control and Prevention. *Vaccines and Immunizations: Rotavirus Vaccine (RotaShield®) and Intussusception*, 2014. Available at: http://www.cdc.gov/vaccines/vpd-vac/rotavirus/vac-rotashield-historical.htm [Accessed on 11 June 2015].
40. Niu MT, Erwin DE, Braun MM. Data mining in the US Vaccine Adverse Event Reporting System (VAERS): early detection of intussusception and other events after rotavirus vaccination. *Vaccine* 2001; **19**:4627–4634.
41. LeBlond R, Brown D, DeGowin R. *DeGowin's Diagnostic Examination* (9th ed). McGraw-Hill: New York, 2009.
42. Yih WK, Lieu TA, Kulldorff M, Martin D, McMahill-Walraven CN, Platt R, Selvam N, Selvan M, Lee GM, Nguyen M. Intussusception risk after rotavirus vaccination in U.S. infants. *New England Journal of Medicine* 2014; **370**:503–512.
43. Ball R, Botsis T. Can network analysis improve pattern recognition among adverse events following immunization reported to VAERS. *Clinical Pharmacology & Therapeutics* 2011; **90**:271–278.
44. Zaversnik M, Batagelj V. *Islands*. Slides from Sunbelt XXIV, Portoroz, Slovenia, May 2004,12–16. Available at: http://vlado.fmf.uni-lj.si/pub/networks/doc/sunbelt/islands.pdf [Accessed on 11 June 2015].
45. Botsis T, Scott J, Goud R, Toman P, Sutherland A, Ball R. Novel algorithms for improved pattern recognition using the US FDA adverse event network analyzer. *Studies in Health Technology and Informatics* 2014; **205**:1178–1182.
46. Best practices recommendations from the Clinical Trials Safety Graphics Working Group. Available at: https://www.ctspedia.org/CTSpedia/BestPractices [Accessed on 3 September 2013].