A Picture Tells 1000 Words (but Most Results Graphs Do Not)

21 Alternatives to Simple Bar and Line Graphs

Jay Hertel, PhD, ATC[a,b,*]

KEYWORDS

• Data visualization • Graphs • Plots • Scientific writing

KEY POINTS

- Bar graphs and line graphs that plot group means and SDs do not provide readers with a thorough understanding of the distribution of individual participant measures.
- Investigators should consider alternatives to bar graphs and line graphs, such as dot plots or box and whisker plots, to visualize individual data points for studies with smaller sample sizes, and violin plots, to display full data distributions in studies with large sample sizes.
- Novel forms of graphs that can illustrate magnitudes of difference, strength of relationships, or multivariate relationships between measures should be considered when presenting research results.

INTRODUCTION

The adage, "a picture tells 1000 words," is often used by experienced investigators when mentoring students, residents, fellows, and other junior colleagues on the intricacies of scientific writing. Graphical representation of research results is often a more effective way to convey findings than text or tables. One wise scholar once told me the ultimate results section of an original research article should contain just 3 words: "see Figure 1." Alas, many investigators default to simple bar graphs or line graphs that are easy to make in common software packages but often have shortcomings when it comes to thoroughly illustrating research findings. When this happens, results figures may not be "worth 1000 words" and, more critically, they may not be a holistic visual representation of the study results.

Disclosure Statement: The author has no conflicts to disclose.

[a] Department of Kinesiology, University of Virginia, PO Box 400407, Charlottesville, VA 22904-4407, USA; [b] Department of Orthopaedic Surgery, University of Virginia, PO Box 800159, Charlottesville, VA 22908, USA

* Department of Kinesiology, University of Virginia, PO Box 400407, Charlottesville, VA 22904-4407.

E-mail address: Jhertel@virginia.edu

Clin Sports Med 37 (2018) 441–462

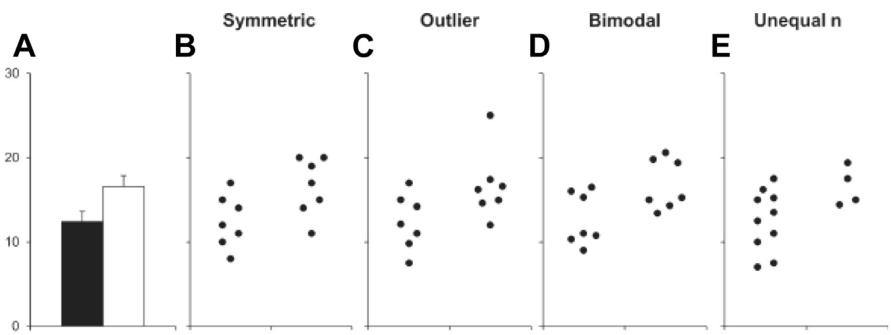
https://doi.org/10.1016/j.csm.2018.04.001 sportsmed.theclinics.com

0278-5919/18/

The primary criticism of bar graphs and line graphs is that drastically different data sets can produce identical mean and SD (or SE) values. This phenomenon was first described by Anscombe[1] in 1973 and has more recently been championed in the life sciences by Weissgerber and colleagues.[2] The primary concern is that differences in group means may be driven by large differences from a small subset of research participants rather than by consistent differences across a majority of participants (Fig. 1). A related concern is that the depiction of data distribution with the group SD (or SE) may be a misleading representation of the distribution of a data set. These concerns have led to a call for investigators to explore alternative ways of illustrating research results with a particular emphasis on graphing the values obtained from individual participants in an effort to allow readers to fully comprehend relationships and trends in a data set.[2–11]

Another concern is that graphs of single, or a select handful of, outcome measures fail to describe the multifactorial relationships that often exist between variables.[7] Investigators frequently limit graphs to only 1 or 2 axes or dimensions, thus placing constraints on how data and relationships can be illustrated and interpreted. Advances in data visualization techniques should be used by investigators in an effort to best represent their research findings to readers. Clinicians and researchers are constantly combing the literature for novel developments in clinical and laboratory techniques; likewise, advances should be sought in methods to visualize research results.

The aim of this article is neither to provide a treatise on statistical distributions and analysis techniques nor to provide a tutorial on the step-by-step procedures of how to construct different types of graphs in specific software programs. Instead, the aim is to provide readers with a (nonexhaustive) set of alternatives to simple bar graphs and line graphs in an effort to spur thought and inspiration about the optimal way to illustrate results.



Test	P value			
t test: Equal var.	0.035	0.050	0.026	0.063
t test: Unequal var.	0.035	0.050	0.026	0.035
Wilcoxon	0.054	0.073	0.128	0.103

Fig. 1. Schematic of 4 data sets with nearly identical group means and SDs (*A*) but very different distributions (*B–E*). The dot plots provide readers with more information on trends in the data set than the bar graph. var, variances. (*From* Weissgerber TL, Milic NM, Winham SJ, et al. Beyond bar and line graphs: time for a new data presentation paradigm. PLoS Biol 2015;13(4):e1002128; with permission.)

ILLUSTRATING INDIVIDUAL PARTICIPANT MEASURES AND GROUP DISTRIBUTIONS

Univariate Scatterplot

Univariate scatterplots, also called dot plots, allow for each subject's measure of a single outcome to be illustrated (see **Fig. 1**B–E). The array of data for each group can be augmented with additional symbols for measures of central tendency, such as mean or median, and variability estimates, such as SD, SE, interquartile range, or CI. A univariate scatterplot can also be transposed on a traditional bar graph. The advantage of this type of graph is that it illustrates all data points and it is particularly useful for data sets with smaller sample sizes.[2]

Box and Whiskers Plot

An extension of the univariate scatter plot is the box and whiskers plot (**Fig. 2**), in which all subjects' data points are graphed vertically (parallel to the Y axis) and a box is formed so that the top and bottom depict the borders of the interquartile range (or other estimate of variability).[12] This box is typically intersected by a horizontal line representing the group median or, on some occasions, the mean. Often, whiskers extending from the top and bottom of the box plot extend to the maximum and minimum values in the data set. The advantage of the box and whiskers plot is that all data points are displayed along with a measure of central tendency and an estimate of variability. Data from different groups or time points can be shown on the same graph to allow for easy visual comparisons of the magnitude and dispersion. A limitation of box plot is that although the height of the box has tangible meaning (variability estimate) the width of the box does not.

Violin Plot

A violin plot builds on the limitation (discussed previously) of the box plot, namely ascribing meaning to both the height and width of the geometric figure that is graphed (**Fig. 3**). The length of the violin plot represents the range of measures in the data set extending from the minimum to the maximum scores, whereas the width represents

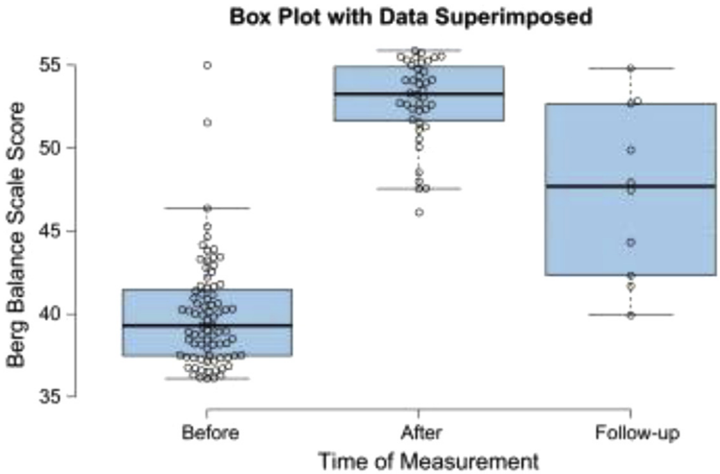


Fig. 2. Box plots provide visualization of the central tendency (median) and distribution (interquartile range and outliers) of a data set. (*From* Nuzzo RL. The box plots alternative for visualizing quantitative data. PM R 2016;8:269; with permission.)

the distribution of individual scores throughout the range. The widest point on a violin plot represents the mode. In the simplest sense, the right and left sides of the shape are mirror images of the histogram of the distribution of individual scores from the sample. Some investigators, however, choose to graph the probability density function rather than the actual sample distribution.[8] Additional encoding may be made within the graph to indicate values of central tendency and variability estimates. Violin plots are particularly useful for illustrating data sets that are large or have non-normal distributions. For example, a data set with a bimodal distribution is readily apparent in a violin plot but not in a box plot.

Graphing Individual Change Scores

Another extension of the univariate scatterplot for pre–post designs can be helpful for illustrating the change in measures for each individual subject between 2 time points[2] (**Fig. 4**). This type of graph can be useful for seeing whether all, or most, subjects had a consistent direction and magnitude of change or whether there is a subset of participants who demonstrated substantial changes (responders) and other subsets of participants who may have demonstrated minimal changes (nonresponders) or changes in the opposite direction. By plotting individual changes over time, this type of graphing does for the visualization of group mean differences what the univariate scatterplot does for group means.

Bland-Altman Plot

A Bland-Altman plot is a specific type of scatterplot that is used to visualize the results of studies comparing 2 measures[13,14] (**Fig. 5**). Specific cases could include the comparison of 2 testing devices or procedures to measure the same construct (eg, comparing maximal voluntary isometric contraction force measures using a handheld dynamometer vs an isokinetic dynamometer), comparison of 2 raters in an assessment of interrater reliability, or comparing repeated measures using the same measurement technique in a test-retest design. The X axis of a Bland-Altman plot represents the mean of the 2 measures taken for each participant, whereas the Y axis represents the arithmetical difference between the 2 measures. Each participant's scores are plotted on the graph yielding visualization of how similar the 2 measurement techniques are (Y axis) across the range of measurement values (X axis). In addition to the X axis, 3 horizontal lines are plotted: the middle line represents the mean difference between the 2 measurements (ideal value is 0) whereas the upper

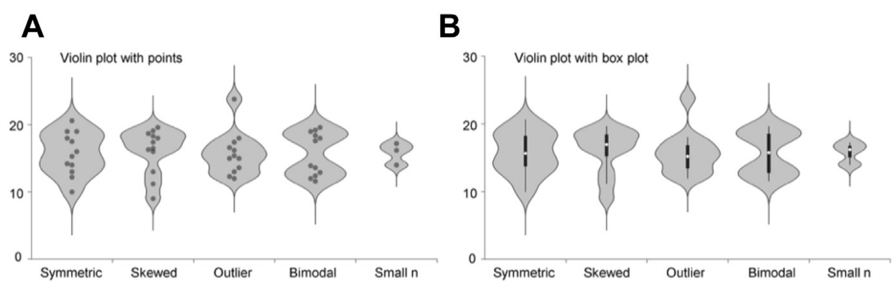


Fig. 3. Violin plots provide visualization of the distribution of a data set and can include presentation of individual data points (*A*) or components of a box and whiskers plot (*B*). (*From* Weissgerber TL, Savic M, Winham SJ, et al. Data visualization, bar naked: a free tool for creating interactive graphics. J Biol Chem 2017;292:20594; with permission.)

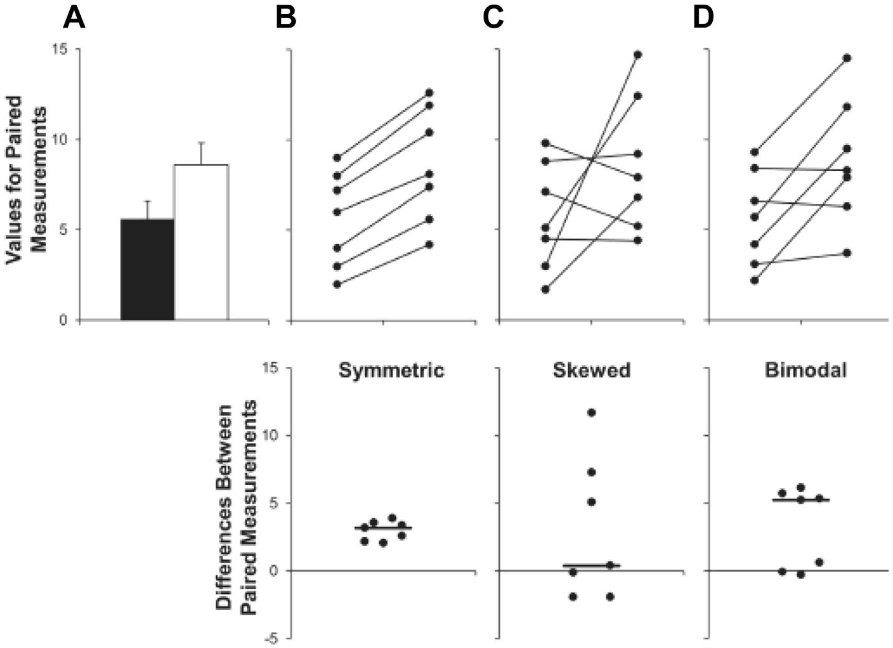


Fig. 4. Visualizing the pre–post change scores of each participant can provide readers with more information about the magnitude and direction of individual change scores than do group metrics. Note that pre–post group changes shown in the bar graph (*A*) could be due to very different pre–post changes in individual measures (*B–D*). (*From* Weissgerber TL, Milic NM, Winham SJ, et al. Beyond bar and line graphs: time for a new data presentation paradigm. PLoS Biol 2015;13(4):e1002128; with permission.)

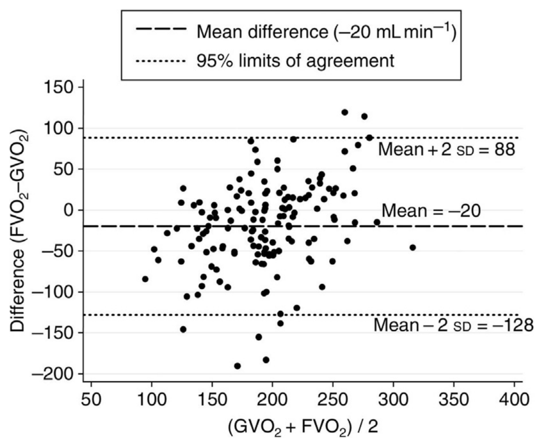


Fig. 5. Bland-Altman plot illustrating the relationship of 2 measurement techniques to assess oxygen consumption: inspired gas analysis (GVO_2) and the reverse Fick method (FVO_2) based on arterial and mixed venous blood gas analysis, respectively. The mean of each pair of measures is plotted on the X axis and the difference between each pair of measures is plotted on the Y axis. (*From* Myles PS, Cui J. Using the Bland–Altman method to measure agreement with repeated measures. Br J Anaesth 2007;99:310; with permission.)

and lower lines represent the 95% limits of agreement between the 2 measurement techniques.[13,14]

ILLUSTRATING PROPORTIONS

Stacked Bar Graph

Although bar graphs have limitations when it comes to illustrating group means and variability estimates, they are beneficial for illustrating frequencies of count data or proportions. Stacked bar graphs are particularly well suited for displaying proportion data when there are a small number of categories (fewer than 4 or 5) within the whole (Fig. 6). An example is expressing the proportion of subgroup members within an entire sample of study participants (eg, the proportion of freshmen, sophomores, juniors, and seniors within a school-based data set). Stacked bar graphs also can be used to express proportions of multiple measures quantified on a continuous scale (eg, the proportion of muscle volume for each of the 4 quadriceps muscles in making up the volume of the entire muscle group), although a limitation of this approach is that it is difficult to illustrate any variability estimates.

Donut Chart

A donut chart is essentially a stacked bar graph that has been converted to a circle where the circumference of the circle, rather than the height of the bar, represents 100% of the sample. The proportion of data in each category is presented in a corresponding portion of the circle's circumference (Fig. 7). A drawback of donut charts is that their usefulness is limited to displaying the results of a single set of measures; thus, they are not useful in showing comparisons across multiple time points.

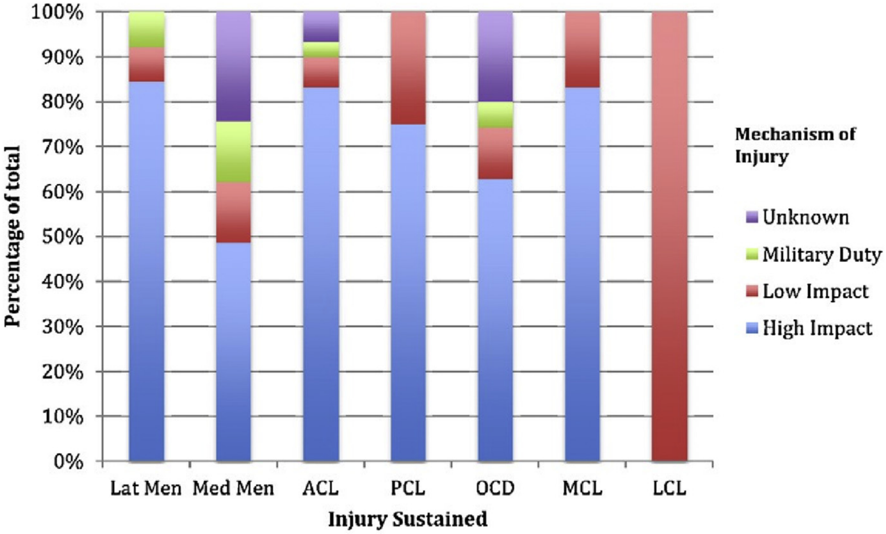


Fig. 6. Stacked bar graph demonstrating proportional data for the mechanism of different knee injuries according to confirmed injury on MRI. ACL, anterior cruciate ligament; Lat, lateral; LCL, lateral collateral ligament; MCL, medical colateral ligament; med, medial; men, meniscal; OCD, osteochondral defect; PCL, posterior cruciate ligament. (*From* Bell D, Wood A, Wrigley S, et al. Assessing the benefit of multidisciplinary assessment centre in a military population sustaining knee injury. J Arthroscopy Joint Surg 2015;2:109; with permission.)

Mosaic Plot

Illustration of the frequency in the outcomes of 2 or more categorical variables may be performed with a mosaic plot, also called a spineplot.[15] Mosaic plots are essentially a pictorial representation of a contingency table. Each combination of categories receives its own bin in which the area is proportional to the frequency of the total sample that lies within it (**Fig. 8**). Unlike bar graphs, where all bars are of the same width (but different height), the width and height of the different bins both can vary to represent smaller or larger proportions.

ILLUSTRATING MAGNITUDE OF DIFFERENCES AND RELATIONSHIPS

Forest Plot

Most often associated with meta-analyses, forest plots are effective at illustrating the magnitude of group differences identified across multiple comparisons. Each comparison represents a "tree," whereas the entire graph embodies the "forest" of related results. The X axis reflects a continuous scale representing the measure of magnitude that could be in the unit of measurement, such as group mean or group mean difference values, or on a unitless scale, such as effect size or odds ratio (**Fig. 9**). For each comparison reported, the point estimate and corresponding variability estimate, usually 95% CI, are plotted on the graph. The results of each comparison are displayed in series from top to bottom as the Y axis is not a numeric scale.

For meta-analyses, the size of the point estimate symbol is often manipulated to reflect the number of participants evaluated in a particular study with larger shapes equating to larger sample sizes. Another unique feature of a forest plot in the context of a meta-analysis is the graphing of the pooled estimate at the bottom of the series of scores (closest to the X axis). The pooled estimate is typically shown as a diamond whose height corresponds to the magnitude of the point estimate and width reflects the CI.

Heat Map

Reporting the results of a study that has a large number of dependent variables can be a challenge for even the most experienced author. Although presenting results in tables is an option, at a certain point the volume of tabular data can become unwieldy when dealing with numerous outcome measures. One approach to data visualization that has gained considerable popularity in recent years is the use of heat maps to illustrate the magnitude of results. In its simplest form, a heat map is a results table that replaces numbers with colors (**Fig. 10**). The spectrum of colors used corresponds to a scale of the numeric values represented. For example, blue, yellow, orange, and red could be operationally defined as representing trivial, small, moderate, and large effect sizes, respectively.

Heat maps are not limited to being color-coded tables. Everyone is likely familiar with weather maps that illustrate the expected temperatures in different geographic areas, where the ends of the temperature spectrum are red equating with hot temperatures and blue with cold temperatures. Similarly, in medical research, heat maps can be used to illustrate the magnitude of measures gathered from different anatomic regions (**Fig. 11**). Such visual representations of results are likely more intuitive to informed readers than sorting through myriad tabular results.

Speedometer Graph

Another means of visualizing the magnitude of results is a speedometer graph, also known as a gauge graph. The scale of this graph is typically a semicircle that is

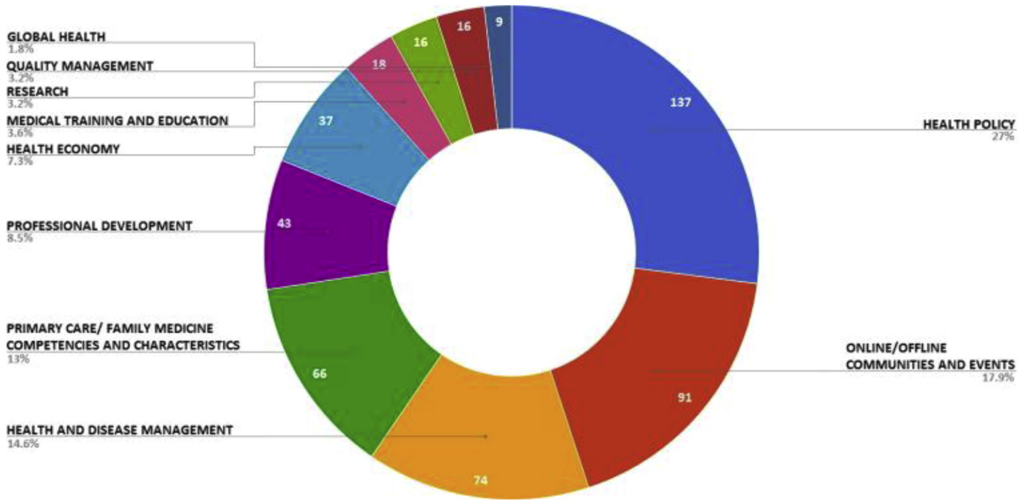


Fig. 7. Donut chart illustrating the proportion of health care–related hashtag themes in the realm of primary care and family medicine from a sample of 500 Twitter posts. The absolute frequency and the percentage of each theme are presented (note that 1.7%, n = 7, of the sample were not coded because they did not match any of the identified themes). (*From* Pinho-Costa L, Yakubu K, Hoedebecke K, et al. Healthcare hashtag index development: identifying global impact in social media. J Biomed Inform 2016;63:395; with permission.)

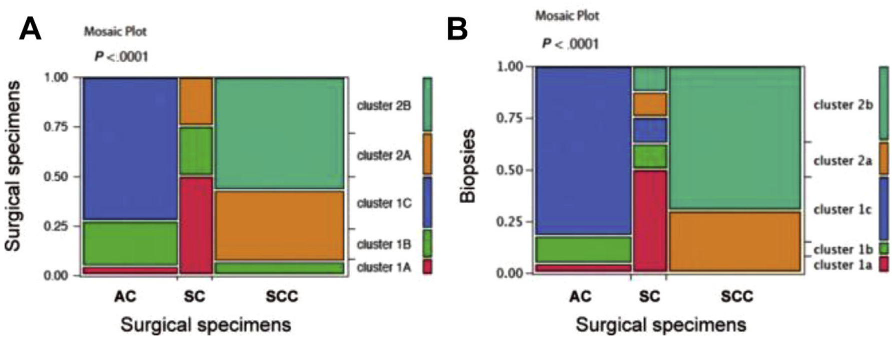


Fig. 8. Mosaic plots illustrating the proportion of different tumor subsets as identified via surgical specimens (*A*) and biopsies (*B*). AC, adenocarcinoma; SC, sarcomatoid carcinoma; SCC, squamous cell carcinoma. (*From* Pelosi G, Rossi G, Bianchi F, et al. Immunhistochemistry by means of widely agreed-upon markers (cytokeratins 5/6 and 7, p63, thyroid transcription factor-1, and vimentin) on small biopsies of non-small cell lung cancer effectively parallels the corresponding profiling and eventual diagnoses on surgical specimens. J Thorac Oncol 2011;6:1044; with permission.)

then labeled with appropriate values. A needle is often used to point to the value of the measure being graphed. A limitation of this approach is that only 1 value may be coherently graphed at a time.

A variation of the speedometer graph may be used to graph several correlation coefficients simultaneously (Fig. 12). This may be useful when comparing an entire

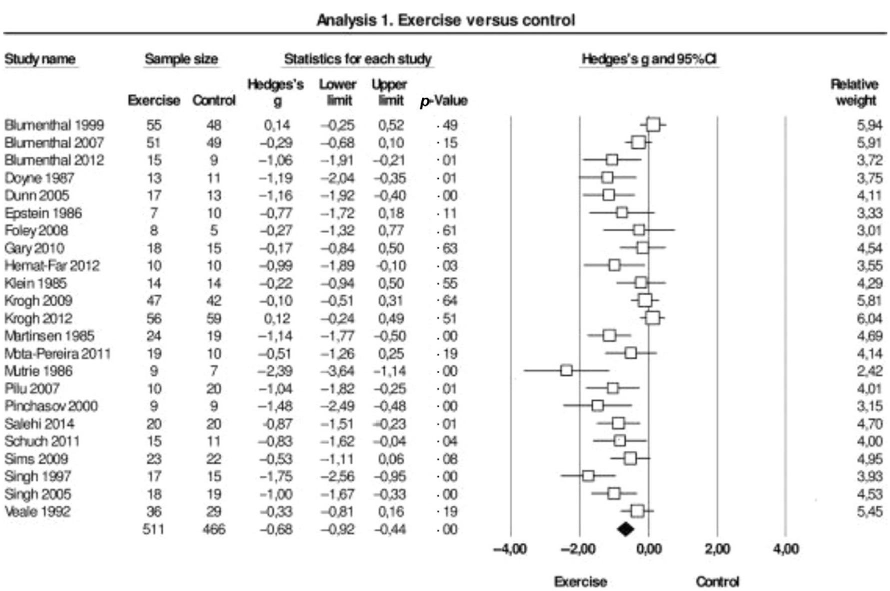


Fig. 9. Forest plot of meta-analysis results evaluating the effects of exercise on the reduction of symptoms related to depression. The open squares represent the effect size point estimate for each of the included studies and the black diamond represents the pooled effect size across all included studies. (*From* Kvam S, Kleppe CL, Nordhus IH, et al. Exercise as a treatment for depression: a meta-analysis. J Affect Disord 2016;202:75; with permission.)

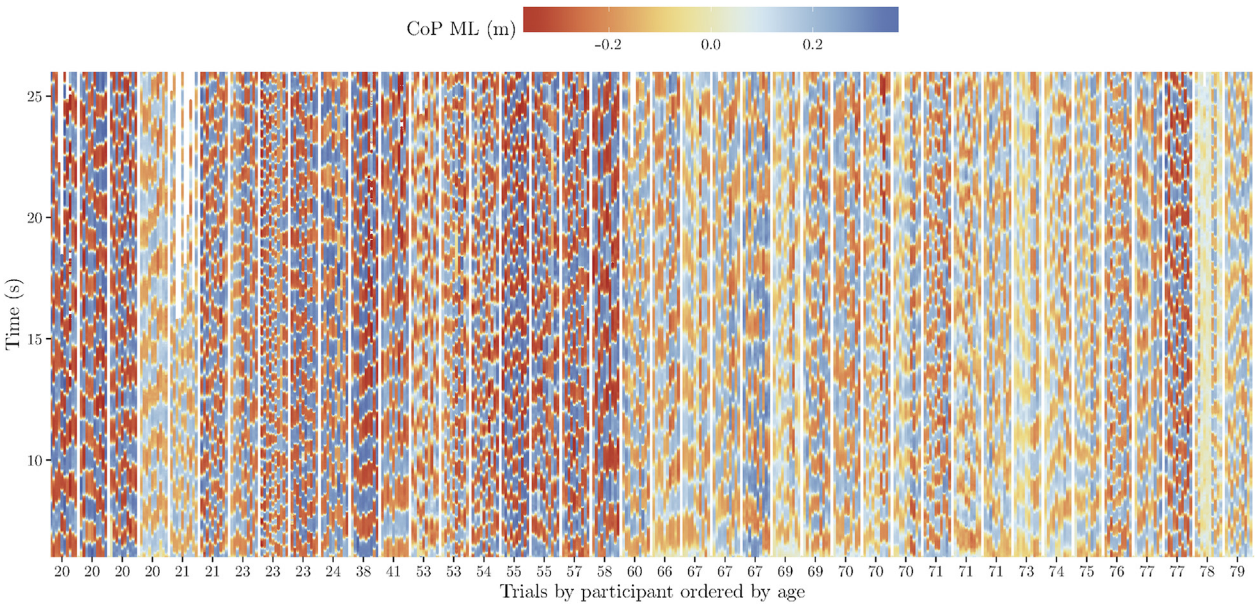


Fig. 10. Heat map of 400 stabilograms of mediolateral position of the center of pressure (CoP ML) during a balance test. The horizontal axis represents the trials per participant, with participants ordered by age. The vertical axis represents time during the trial, ranging from 6 to 26 seconds. The color scale for the position of CoP ML is at the top of the figure. (*From* Soancatl Aguilar V, van den Gronde J, Lamoth C, et al. Visual data exploration for balance quantification in real-time during exergaming. PLoS One 2017;12(1):e0170906; with permission.)

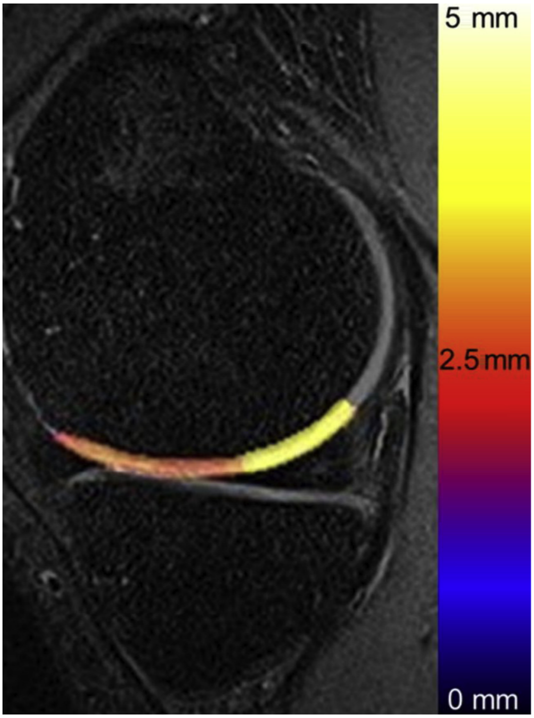


Fig. 11. Heat map depicting cartilage thickness on the medial femoral condyle. (*From* Schmitz RJ, Wang HM, Polprasert DR, et al. Evaluation of knee cartilage thickness: a comparison between ultrasound and magnetic resonance imaging methods. Knee 2017;24:219; with permission.)

matrix of dependent variables to each other. The midpoint of the gauge is labeled with a value of 0 (no correlation), whereas the far right end of the gauge is labeled with a value of +1.0 (perfect positive correlation), and the far left end is labeled with a value of −1.0 (perfect negative correlation). Symbols representing each pair of bivariate correlations are then placed on the graph to indicate the appropriate correlation coefficient point estimate. The basis for this type of figure is the Taylor diagram,[16] a more complex graph that is used to illustrate the correlation coefficient, root mean square error, and the SD between a modeled system and observed behavior. This described variation uses only the correlation coefficient portion of the Taylor diagram.

ILLUSTRATING TWO MEASURES SIMULTANEOUSLY

Dual Y-axis Graph

Investigators are often faced with the challenge of illustrating the results of more than 1 dependent measure within the context of a research study. Although each measure could be represented on a separate graph, this may cloud the relationships between the changes in multiple measures and is an inefficient use of space. One potential solution is the use of a graph with dual Y axes in which the Y axis on the left side of the graph is scaled for the primary outcome and the Y axis on the right side of the graph is scaled for the secondary outcome (Fig. 13). The X axis often consists of a time series. The 2 outcomes may be graphed in 1 of these combinations: (1) 1 as a bar graph and

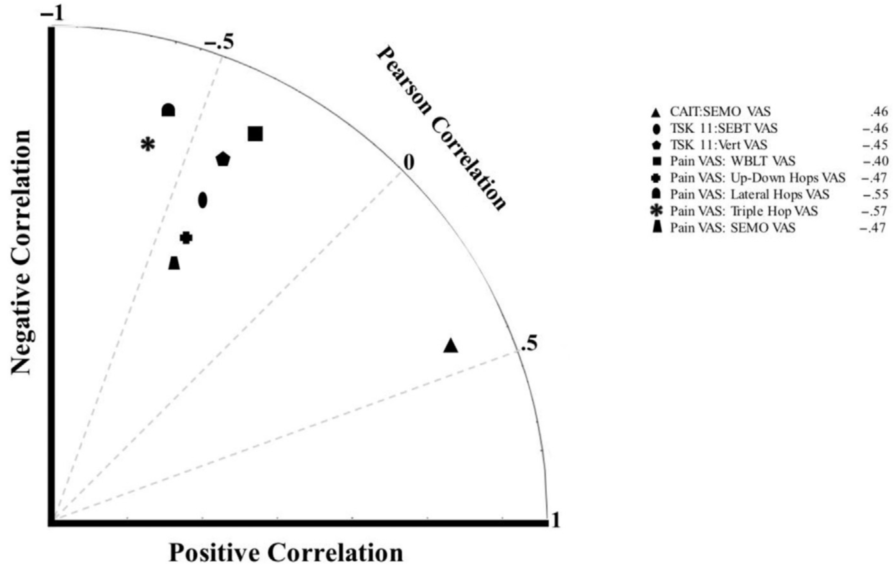


Fig. 12. A modified speedometer graph illustrating the magnitude of Pearson correlation coefficients from several correlation analyses comparing scores of self-reported questionnaires and perceived confidence scores when performing functional tests among a group of high school athletes with a history of ankle sprain. Note that the curved scale ranges from negative 1.0 at the top left to positive 1.0 at the bottom right. CAIT, Cumberland ankle instability tool; SEBT, star excursion balance test; SEMO, southeast Missouri agility test; TSK 11, Tampa scale of kinesiophobia 11; VAS, visual analog scale; Vert, vertical jump; WBLT, weight-bearing lunge test. Note: All reported r values are statistically significant; $P<.05$. This figure is derived from unpublished data (Revay O. Corbett, 2018) from the Exercise & Sport Injury Lab at the University of Virginia. (*Courtesy of* Revay O. Corbett, MS, ATC; with permission.)

the other as a line graph, (2) both as line graphs, and (3) both as bar graphs. Clear labeling of the axes and outcome measures is key to ensuring reader comprehension of these graphs.

Angle-Angle Plot

In musculoskeletal biomechanics research, it is common to simultaneously measure motion at more than 1 joint or in more than 1 plane at a single joint. An alternative to illustrating 2 streams of kinematic data independently is to plot them concurrently on an angle-angle plot (**Fig. 14**). One plane of motion is plotted on the X axis with the other plane on the Y axis. For each data sample, the kinematic values are plotted as in a scatterplot, and then a line is drawn to connect adjacent data points. Angle-angle plots are particularly effective for illustrating the coupling behavior of joint motions. A limitation of angle-angle plots is that there is not a time series on either axis, thus making it essential that the investigator clearly label important temporal events (eg, heel strike and toe-off during gait) on the graph.

Vector Graph

An effective manner of graphing simultaneous motion, or force, in 2 planes between 2 discrete time points is the use of vectors (**Fig. 15**). As in the angle-angle plot, 1

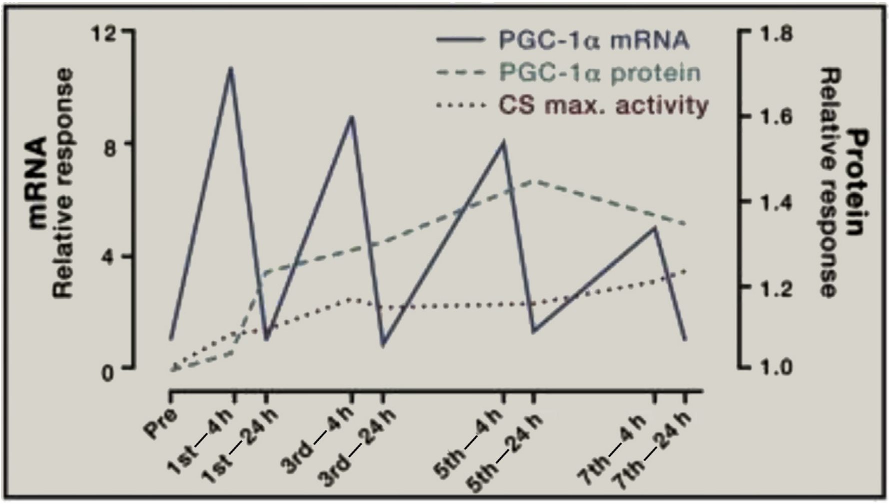


Fig. 13. A dual Y-axis graph with the scale for the relative response of mRNA (on the left Y axis) and protein (on the right Y axis) to 7 sessions of high-intensity interval training during a 2-week exercise intervention. Skeletal muscle biopsies from the vastus lateralis were obtained 4 hours and 24 hours after the first, third, fifth, and seventh training sessions. CS, citrate synthase; max, maximum; PGC-1α, peroxisome proliferator-activated receptor γ coactivator α. (*From* Hawley JA, Hargreaves M, Joyner MJ, et al. Integrative biology of exercise. Cell 2014;159(4):743; with permission.)

plane of motion is plotted on the X axis whereas a second plane of motion is plotted on the Y axis. Using a starting time point at the origin (0, 0 point) of the graph, the ending time point is plotted and the resultant vector is calculated. The length of the resultant vector represents the magnitude of coupled motion, whereas the angle of

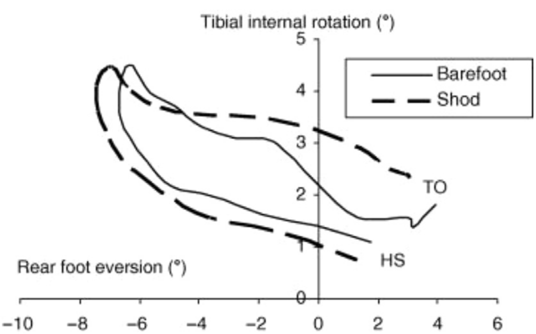


Fig. 14. An angle-angle plot illustrating the coupled motion of rearfoot inversion-eversion on the X axis and tibial internal-external rotation on the Y axis during the stance phase of running. Subjects ran in barefoot and shod conditions. Note that time is not depicted on either axis but the occurrence of heel strike (HS) and toe-off (TO) are clearly labeled on the graph to provide a temporal orientation. (*From* Eslami M, Begon M, Farahpour N, et al. Forefoot–rearfoot coupling patterns and tibial internal rotation during stance phase of barefoot versus shod running. Clin Biomech (Bristol, Avon) 2007;22:77; with permission.)

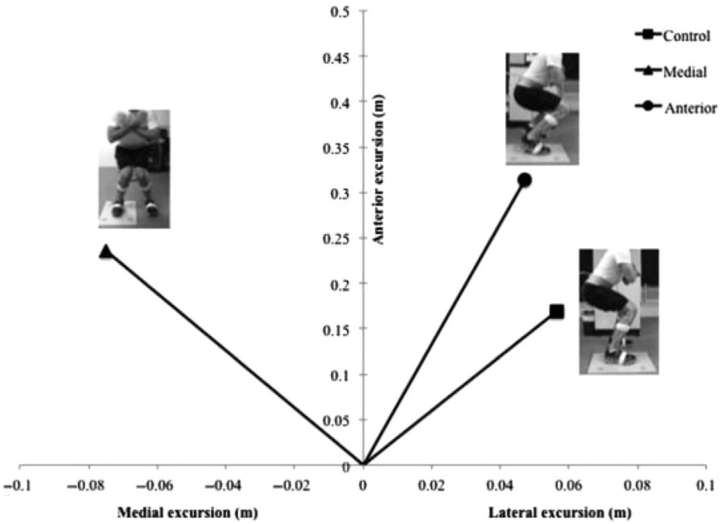


Fig. 15. This vector graph illustrates knee displacements simultaneously in the anterior and medial-lateral directions during squatting exercises performed with 3 different sets of instruction (control, anterior malaligned, and medial malaligned). The 0, 0 coordinate represents the starting position in upright stance whereas the position during maximal squat is represented by the symbols. (*From* Slater LV, Hart JM. The influence of knee alignment on lower extremity kinetics during squats. J Electromyogr Kinesiol 2016;31:98; with permission.)

the resultant vector from the horizontal indicates the ratio of motion between the 2 planes. The most common application of this type of analysis is called *vector coding* and is used to serially quantify the joint coupling behavior across movement tasks, such as gait.[17] It can, however, also be used to clearly graph coupled motion between any 2 discrete time points.

Receiver Operator Curve

In diagnostic accuracy studies, the estimation of sensitivity and specificity of diagnostic tests is central to the research questions being asked. For a diagnostic test that is scored on a continuous or ordinal scale, the establishment of threshold values for whether a test is positive or negative, the use of a receiver operator characteristic (ROC) curve can be helpful (Fig. 16). The ROC graph is constructed with sensitivity on the Y axis with values ranging from 0 to 1.0, whereas 1 minus specificity values on the X axis with values ranging from 0 to 1.0. This orientation places a test with both high sensitivity and high specificity in the upper left (or northwest) corner of the graph. There is also a diagonal line extending from the lower left hand corner to the upper right hand corner of the graph. Combinations of sensitivity and specificity that lie above that line are associated with diagnostic procedures with greater than a 50-50 chance of producing a correct diagnosis, whereas combinations below the line are less than chance.

To establish the best diagnostic threshold value, sensitivity and specificity are calculated at each possible threshold value in a data set. These combinations of values are then plotted on the ROC graph and a line is generated that extends from the lower left hand corner of the graph through each point plotted on the graph before

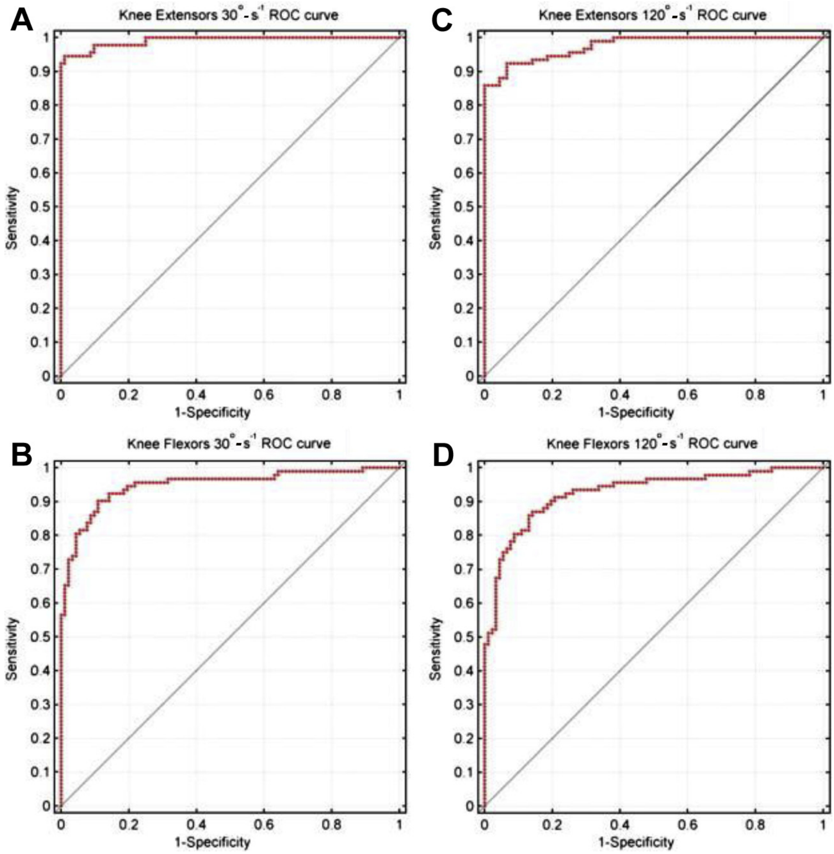


Fig. 16. A set of ROC curves of 4 decision rule models aimed at identifying feigned performance during maximal isokinetic strength testing. Note that the highest combination of sensitivity and specificity and the associated area under the curve values are found in descending order in graphs *A–D*. (*From* Almosnino S, Stevenson JM, Day AG, et al. Discriminating between maximal and feigned isokinetic knee musculature performance using waveform similarity measures. Clin Biomech (Bristol, Avon) 2012;27:381; with permission.)

culminating in the upper right hand corner of the graph. The point plotted in the most northwest position on the graph represents the optimal diagnostic threshold value and the area under the ROC curve represents the diagnostic accuracy of that threshold. The ROC curves of different diagnostic procedures may be compared on a single graph.

ILLUSTRATING THREE OR MORE MEASURES SIMULTANEOUSLY

Bubble Chart

A typical scatterplot graphs the scores of 2 measures, 1 on the X axis and the other on the Y axis, of individual subjects. The appearance of the symbols for each subject are typically uniform, although sometimes additional information can be conveyed, such as plotting the data of female subjects with filled circles and male subjects with clear circles. Bubble graphs expand on the simple scatterplot by changing the size of each subject's symbol based on the value of a third measure (the first 2 measures displayed on the X and Y axes) (Fig. 17). A limitation of simple scatterplots is that they only allow for the simultaneous presentation of 2 measures.

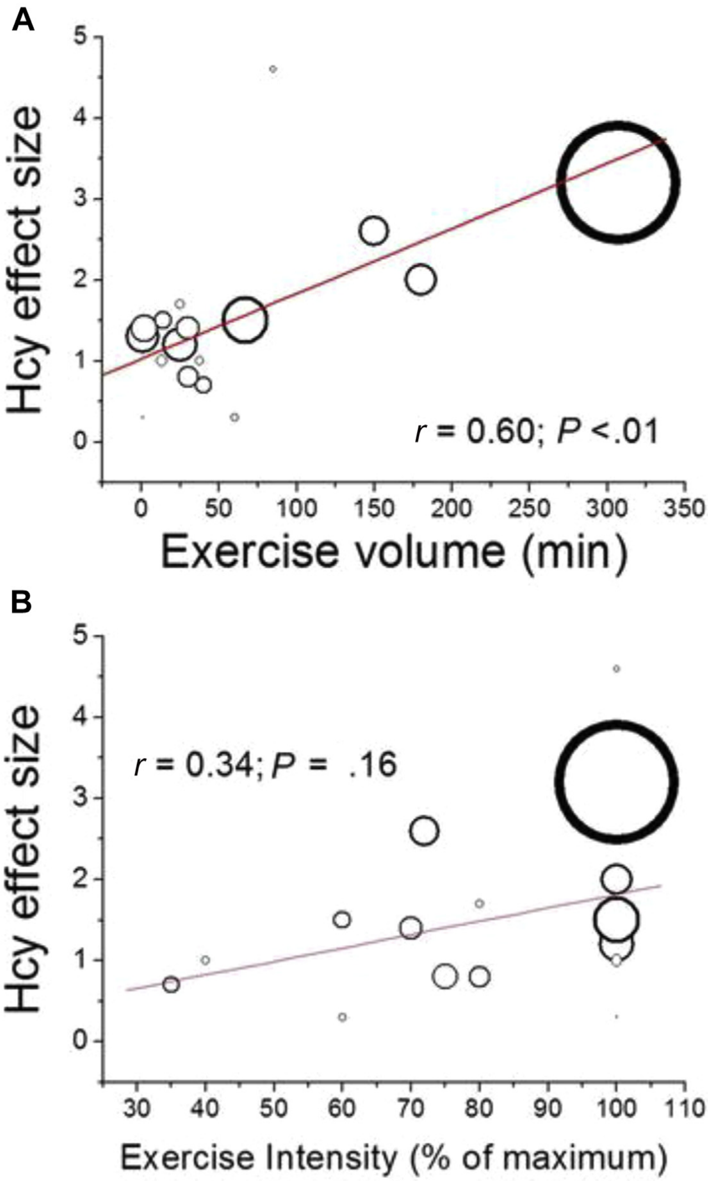


Fig. 17. A pair of bubble charts illustrating the dose-response relationship of exercise volume (*A*) and intensity (*B*) on the effect size changes of homocysteine (Hcy) across 18 separate studies. Each circle represents the results on a specific study and the size of the circle corresponds to the precision of each estimate to the fitted regression line. min, minutes. (*From* Deminice R, Ribeiro DF, Frajacomo FT. The effects of acute exercise and exercise training on plasma homocysteine: a meta-analysis. PLos One 2016;11(3):e0151653; with permission.)

3-D Scatterplot

Another useful way to visualize relationships between 3 different measures is the 3-D scatterplot (**Fig. 18**). The response variable is typically graphed on the Z axis whereas the 2 predictors variables are graphed on the X and Y axes. Software programs used to construct 3-D scatterplots typically allow for users to rotate the cube-shaped graph in an effort to find the best perspective for readers to see the direction and strength of the relationships between the 3 variables. Choosing a single 2-D view from which readers can view the graph on a static page in a printed article can be a challenge for investigators. Interactive graphing tools can counteract this challenge.

Radar Chart

A radar chart, also called a spider chart, may be used to simultaneously visualize the scores of 3 or more continuous or ordinal variables (**Fig. 19**). Each variable is plotted on its own spoke, or radius, extending outward from the center of the chart. The number of variables dictate the angle at which the spokes deviate from each other. Each spoke may have its own measurement scale although it can be confusing to readers when the scales differ. Measures from a single participant or group are plotted on each spoke and a line is then drawn to connect the scores of all measures. Results from additional participants and groups are then added to the chart, allowing for comparisons in the pattern of scores. Consideration should be given to the order in which the various outcome measures are displayed on the graph because differing orders can produce dissimilar shapes and lead to spurious interpretations of results.

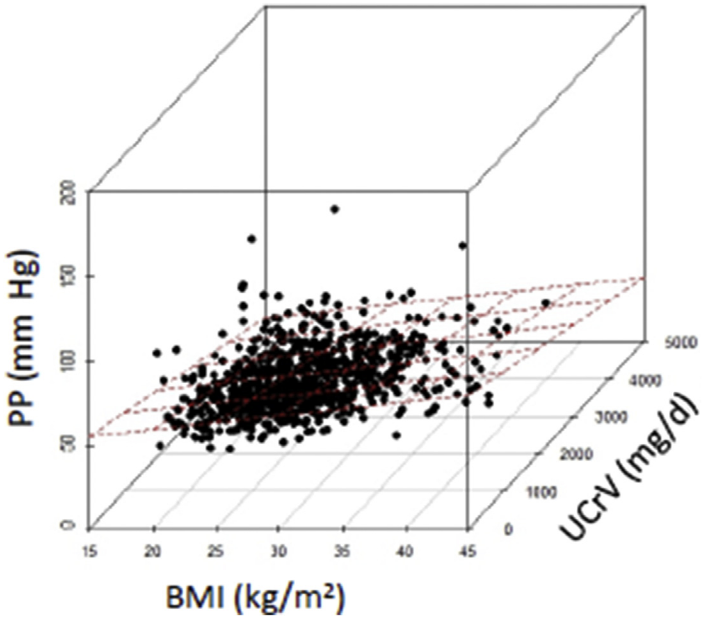


Fig. 18. A 3-D scatterplot illustrating the relationships between pulse pressure (PP), body mass index (BMI), and urinary creatinine excretion rate (UCrV) in 840 patients with chronic kidney disease. The red plane represents the plane of prediction for PP scores from BMI and UCrV scores. (*From* Shah PT, Martin R, Sanabria J, et al. Adiposity predicts pulse pressure in patients with chronic kidney disease: data from the modification of diet in renal disease. J Cardiol Curr Res 2016;7(2):00240; with permission.)

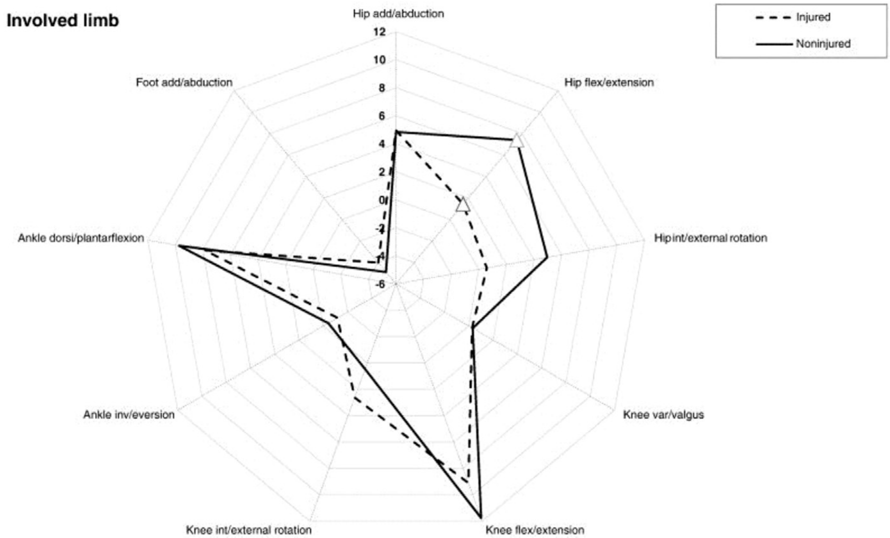


Fig. 19. A radar graph displaying the average joint position of the hip, knee, and ankle during a 20-second unipedal balance task for the involved limb in injured (ankle sprain) and uninjured groups. For each plane of motion, movements are listed in order of positive and negative values with neutral position equating to 0. Δ, statistically significant between-groups difference. (*From* Doherty C, Bleakley C, Hertel J, et al. Postural control strategies during single limb stance following acute lateral ankle sprain. Clin Biomech (Bristol, Avon) 2014;29:646; with permission.)

ILLUSTRATING DATA ACROSS TIME

Graphing Means and CIs Over Time

The visual presentation of time series data can present challenges, especially in regard to graphing variability estimates. One criticism of line graphs is that, similar to bar graphs, visualizing the distribution of individual participant scores within a data set can be difficult. When a time series consists of numerous measures taken at consistent intervals, 1 solution may be graphing 3 lines across the entire time series, with the middle line representative of the group mean and the top and bottom lines indicating the upper and lower boundaries of the associated CI around the mean. When this is done for 2 groups on the same graph, readers can visually assess the magnitude of differences in the group means as well as whether or not the CIs for the 2 groups overlap (Fig. 20). Some investigators specifically highlight the time epochs where the CIs do not overlap because these regions are often interpreted as significantly different from each other.

Time to Event Graph

Dichotomous outcomes, such as injury prevention (injured, not injured) or return to play after injury (returned, has not returned) are time series data typically assessed with survival analysis. Visualization of such data may be performed with Kaplan-Meier curves, where the follow-up time is expressed on the X axis and the proportion of participants who have or have not experienced the outcome of interest is displayed on the Y axis (Fig. 21). Group results showing the decline in the proportion of participants who not yet experienced the outcome of interest are plotted over time. The curves for 2 or more groups may be illustrated on the same graph for comparison.

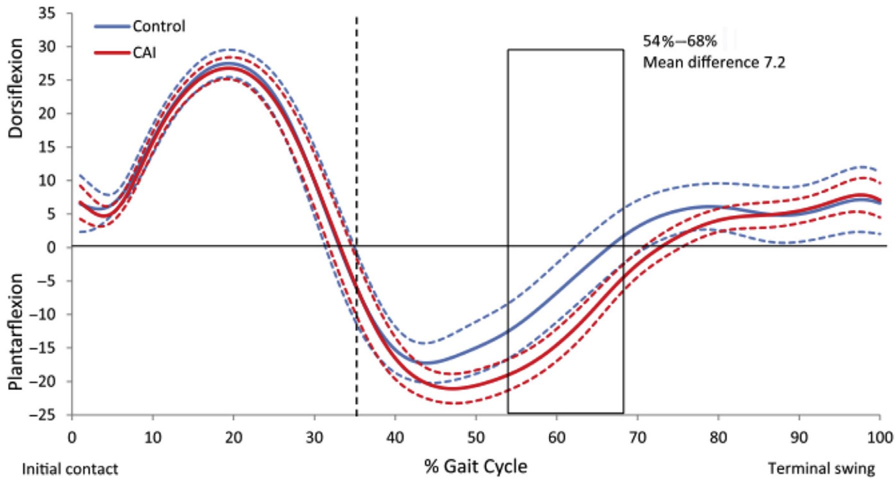


Fig. 20. A figure graphing the mean and 95% CI across the entire jogging gait cycle for a group with chronic ankle instability (CAI) and a healthy control group. The time points where the CIs for the 2 groups do not overlap are considered to be significantly different and are outlined in the rectangle. (*From* Chinn L, Dicharry J, Hertel J. Ankle kinematics of individuals with chronic ankle instability while walking and jogging on a treadmill in shoes. Phys Ther Sport 2013;14:236; with permission.)

Stacked Area Graph

An area graph is essentially a line graph with the area beneath the line filled in with color down to the X axis. The X axis is typically a time series. Area graphs may be a useful way to illustrate changes in a single measure over time. A limitation of area graphs is the lack of presentation of variability estimates.

If there is more than 1 category of data to be displayed on the same time series, a stacked area graph may be constructed to visualize the how the various categories change over time (Fig. 22). One limitation of stacked area graphs is that although the category shown on the bottom of the graph starts at the bottom of the measurement scale shown on the Y axis (usually 0); this first category's data are intuitive and easy to interpret. The other categories' scores, however, are stacked on top on another group's data and do not start at 0. Thus, stacked bar graphs may be better used for illustrating comparisons of proportional differences over time rather than comparison of absolute values in each category at each time point.

CAVEATS

The list of graph types described in this article are in no way meant to be all-inclusive. Other types of graphs may be more appropriate for illustrating a given set of results than those options discussed. In some cases, the best option may be a simple bar graph or line graph and investigators should be comfortable making that decision. This article is not a call for a permanent ban of all bar graphs and line graphs.

Readers must also be cognizant that the recommendations provided may become outdated with advances in data analytics and visualization. All the graphs described are static in nature and meant for the printed page. As more journals continue to move to online-only publications, investigators must recognize that journal articles at some point will move beyond the printed page (or PDF file) and possibilities like

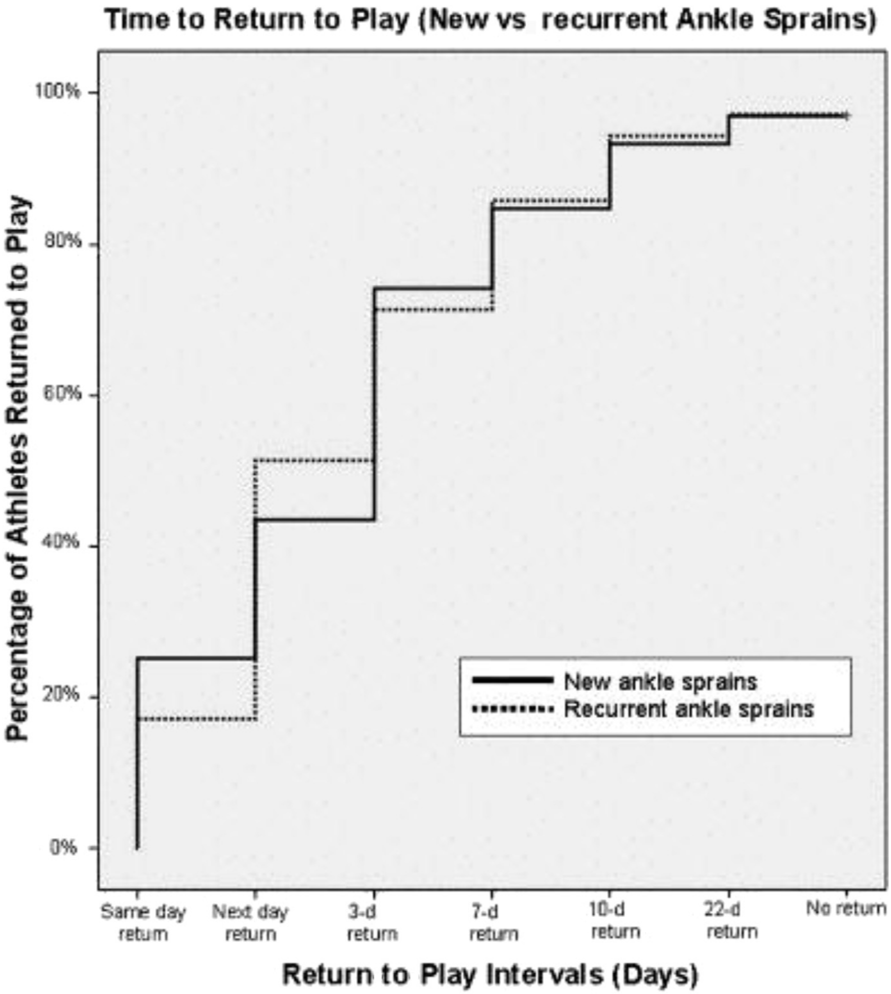


Fig. 21. A time to event graph illustrating the percentage of high school athletes who have returned to play at different time intervals after new versus recurrent ankle sprains. There was no significant difference in the return to play probabilities between these 2 groups. (*From* Medina McKeon JM, Bush HM, Reed A, et al. Return-to-play probabilities following new versus recurrent ankle sprains in high school athletes. J Sci Med Sport 2014;17:26; with permission.)

animated graphics and interactive figures may render many graphing techniques that are now accepted obsolete.[18]

RECOMMENDATIONS

The following recommendations are provided:

- A list of helpful graphing resources may be found in **Table 1**.
- Investigators should consider the importance of designing a signature figure for each research article they publish. Consider not only how this figure will look on the printed page but also how it will look on social media posts or on a slide presented by someone outside the research group.

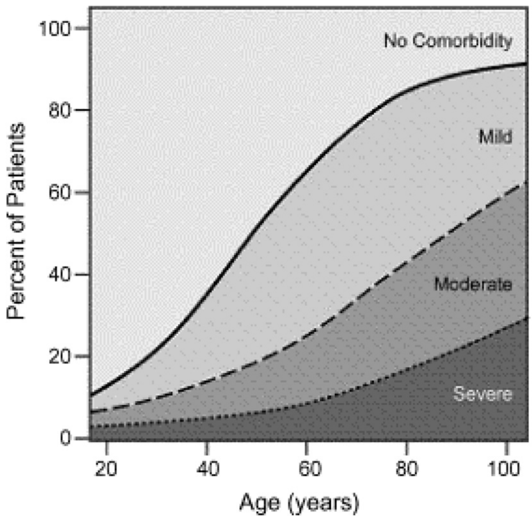


Fig. 22. A stacked area graph illustrating the percentage of 27,506 adult cancer patients with comorbidities across the age spectrum. (*From* Piccirillo JF, Vlahiotis A, Barrett LB, et al. The changing prevalence of comorbidity across the age spectrum. Crit Rev Oncol Hematol 2008;67:126; with permission.)

- Brainstorm multiple ways to graph a set of results before settling on how it will be done.
- Be cognizant of illustrating the results of individual participants and group distributions whenever possible.
- Clearly label each graph, paying particular attention to the labeling of the axes; the size, shape, and color of lines, bars, and other symbols; and how statistical significance and the magnitude of differences or strength of relationships are depicted.

Table 1
Selected resources to aid in making innovative graphs.

Resource	Website
Chart.js	www.chartjs.org
Creative Bloq	www.creativebloq.com/design-tools/data-visualization-712402
Data Hero	www.datahero.com
Matlab	www.mathworks.com
Microsoft Excel	www.support.office.com/en-us/excel
Plotly	www.plot.ly
R	www.r-project.org
RawGraphs	www.rawgraphs.io
SAS	www.sas.com
SPSS	www.spss.com
Tableau	www.tableau.com

Explore the following Web sites to learn more about using software packages and online tools for basic and advanced graphing functions. This list is not meant to be an all-inclusive list of graphing tools.

- Write a clear and informative caption for each graph in a manner that orients readers to the graph and guides them through the most important results.
- Prior to article submission, show graphs and captions to a colleague who is unfamiliar with the study and ask if the results are clearly presented.

REFERENCES

1. Anscombe FJ. Graphs in statistical analysis. Am Stat 1973;27:17–21.
2. Weissgerber TL, Milic NM, Winham SJ, et al. Beyond bar and line graphs: time for a new data presentation paradigm. PLoS Biol 2015;13(4):e1002128.
3. Schriger DL, Cooper RJ. Achieving graphical excellence: suggestions and methods for creating high-quality visual displays of experimental data. Ann Emerg Med 2001;37:75–87.
4. Cooper RJ, Schriger DL, Close RJ. Graphical literacy: the quality of graphs in a large-circulation journal. Ann Emerg Med 2002;40:317–22.
5. Lane DM, Sandor A. Designing better graphs by including distributional information and integrating words, numbers, and images. Psychol Methods 2009;14:239–57.
6. Paniello RC, Neely JG, Rich JT, et al. Practical guide to choosing an appropriate data display. Otolaryngol Head Neck Surg 2011;145:886–94.
7. Sperandei S. The pits and falls of graphical presentation. Biochem Med 2014;24:311–20.
8. Correll M, Gleicher M. Error bars considered harmful: exploring alternate encodings for mean and error. IEEE Trans Vis Comput Graph 2014;20:2142–51.
9. Duke SP, Bancken F, Crowe B, et al. Seeing is believing: good graphic design principles for medical research. Stat Med 2015;34:3040–59.
10. Rice K, Lumley T. Graphics and statistics for cardiology: comparing categorical and continuous variables. Heart 2016;102:349–55.
11. Weissgerber TL, Savic M, Winham SJ, et al. Data visualization, bar naked: a free tool for creating interactive graphics. J Biol Chem 2017;292:20592–8.
12. Nuzzo RL. The box plots alternative for visualizing quantitative data. PM R 2016;8:268–72.
13. Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. Statistician 1983;32:307–17.
14. Giavarina D. Understanding bland altman analysis. Biochem Med 2015;25:141–51.
15. Pelosi G, Rossi G, Bianchi F, et al. Immunhistochemistry by means of widely agreed-upon markers (cytokeratins 5/6 and 7, p63, thyroid transcription factor-1, and vimentin) on small biopsies of non-small cell lung cancer effectively parallels the corresponding profiling and eventual diagnoses on surgical specimens. J Thorac Oncol 2011;6:1039–49.
16. Taylor KE. Summarizing multiple aspects of model performance in a single diagram. J Geophys Res 2001;106:7183–92.
17. Herb CC, Chinn L, Dicharry J, et al. Shank-rearfoot joint coupling with chronic ankle instability. J Appl Biomech 2014;30:366–72.
18. Weissgerber TL, Garovic VD, Savic M, et al. From static to interactive: transforming data visualization to improve transparency. PLoS Biol 2016;14(6):e1002484.