



# Rethinking data analysis in basic biomedical research

Natasa Milic, MD, PhD



# FEATURE



---

CHRISTMAS 2015: THE PUBLICATION GAME

## Rejection of rejection: a novel approach to overcoming barriers to publication

Tired of rejections? **Cath Chapman** and **Tim Slade** offer a simple way to fight back

Cath Chapman *senior research fellow*, Tim Slade *associate professor*

NHMRC Centre of Research Excellence in Mental Health and Substance Use, National Drug and Alcohol Research Centre, University of New South Wales, Sydney, Australia

## Rejection of rejection letter

*[insert university emblem here]*

Dear Professor *[insert name of editor]*

*[Re: MS 2015\_XXXX Insert title of ground-breaking study here]*

Thank you for your rejection of the above manuscript.

Unfortunately we are not able to accept it at this time. As you are probably aware we receive many rejections each year and are simply not able to accept them all. In fact, with increasing pressure on citation rates and fiercely competitive funding structures we typically accept fewer than 30% of the rejections we receive. Please don't take this as a reflection of your work. The standard of some of the rejections we receive is very high.

In terms of the specific factors influencing our decision the failure by Assessor 1 to realise the brilliance of the study was certainly one of them. Simply stating "this study is neither novel nor interesting and does not extend knowledge in this area" is not reason enough. This, coupled with the use of Latin quotes by Assessor 2, rendered an acceptance of your rejection extremely unlikely.

We do wish you and your editorial team every success with your rejections in the future and hope they find safe harbour elsewhere. To this end, may we suggest you send one to *[insert name of rival research group]* for consideration. They accept rejections from some very influential journals.

Please understand that our decision regarding your rejection is final. We have uploaded the final manuscript in its original form, along with the signed copyright transfer form.

We look forward to receiving the proofs and to working with you in the future.

Yours sincerely

Dr *[insert name here]*

*[Insert research group acronym here]*

*[Insert university here]*

*[Insert country here—that is, Australia/New Zealand/small European Country/Canada]*

# Increased quantification of scientific research

- have expanded the scope of applications of statistical methods
- concept of “statistical significance”

In February, 2014, George Cobb, Professor Emeritus of Mathematics and Statistics at Mount Holyoke College, posed these questions to an ASA discussion forum:

- Q: Why do so many colleges and grad schools teach  $p = 0.05$ ?
- A: Because that's still what the scientific community and journal editors use.
- Q: Why do so many people still use  $p = 0.05$ ?
- A: Because that's what they were taught in college or grad school.

# Cobb's concern

- Cobb's concern was a long-worrisome circularity in the sociology of science based on the use of bright lines such as  $P < 0.05$  :
- “We teach it because it's what we do; we do it because it's what we teach.”

# Highly visible discussions over the last few years

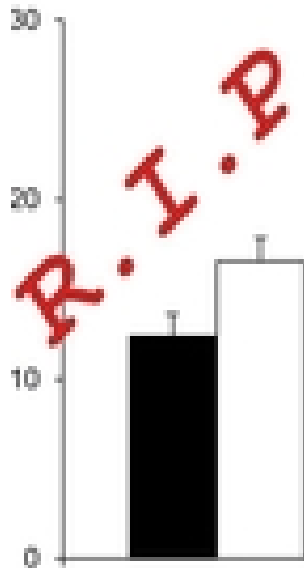
- ScienceNews (Siegfried, 2010): “It’s science’s dirtiest secret: The ‘scientific method’ of testing hypotheses by statistical analysis stands on a flimsy foundation.”
- Phys.org Science News Wire (2013): “numerous deep flaws” in null hypothesis significance testing
- ScienceNews (Siegfried, 2014): “statistical techniques for testing hypotheses...have more flaws than Facebook’s privacy policies.”
- “Simply Statistics” (Leek, 2014): “The problem is not that people use P-values poorly, it is that the vast majority of data analysis is not performed by people properly trained to perform data analysis”

2,104 Save	54 Citation
228,066 View	6,781 Share

PERSPECTIVE

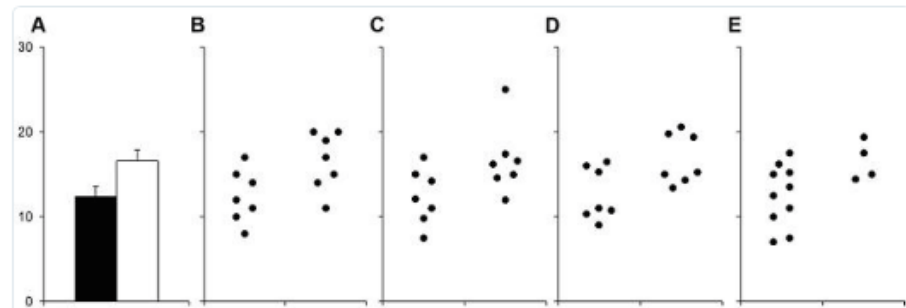
# Beyond Bar and Line Graphs: Time for a New Data Presentation Paradigm

Tracey L. Weissgerber<sup>1\*</sup>, Natasa M. Milic<sup>1,2</sup>, Stacey J. Winham<sup>3</sup>, Vesna D. Garovic<sup>1</sup>



PLOS Biology @PLOSbiology · Apr 23

Bar graphs hide the distribution of continuous data - time to kill them off! #PLOSbiology  
[plos.io/1G1xeRL](https://plos.io/1G1xeRL)





April 2015

NATURE | RESEARCH HIGHLIGHTS: SOCIAL SELECTION



## Bar graphs criticized for misrepresenting data

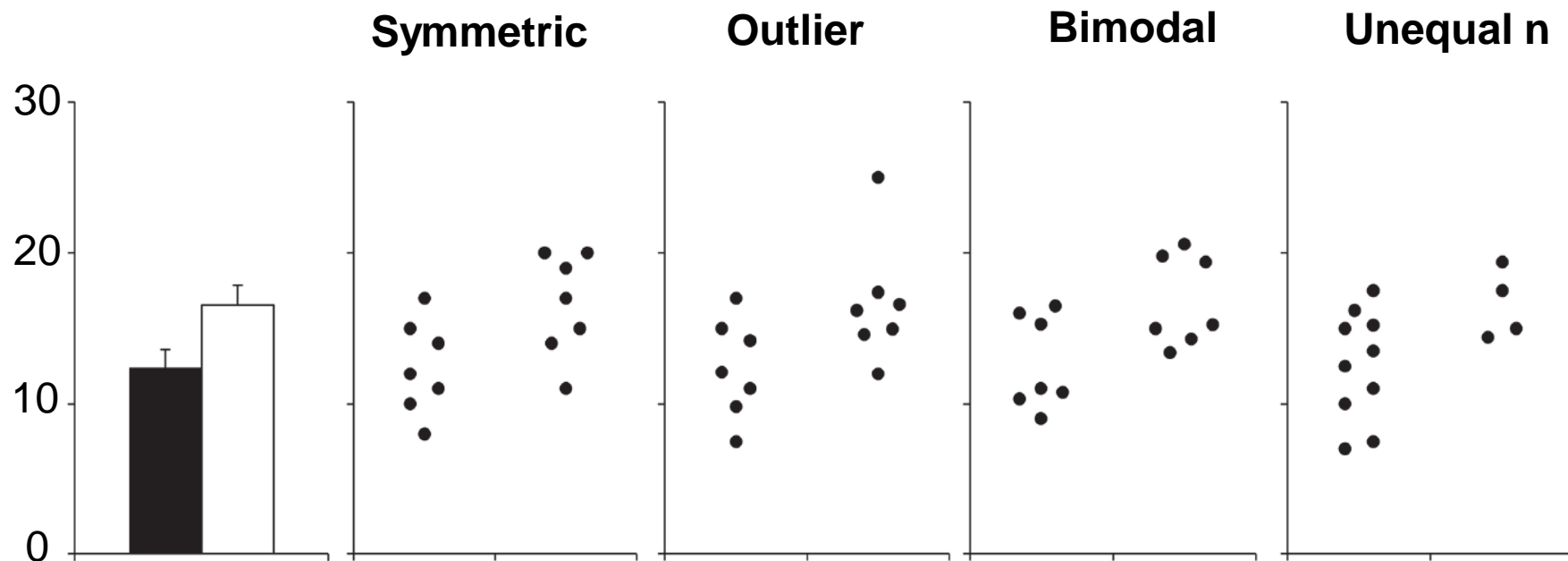
Researchers on Twitter denounce a common way to display data.

- ❖ Researchers at the Mayo Clinic in Rochester, Minnesota
- ❖ looked at more than 600 research articles
- ❖ published in top physiology journals in early 2014
- ❖ and found that bar graphs were used to describe continuous data
- ❖ in more than 85% of the articles
- ❖ They showed that very different data sets can be described by the same bar graph
- ❖ As an alternative, they recommend dotplots showing every data point, especially for studies with small sample sizes

# Key Findings: Systematic Review

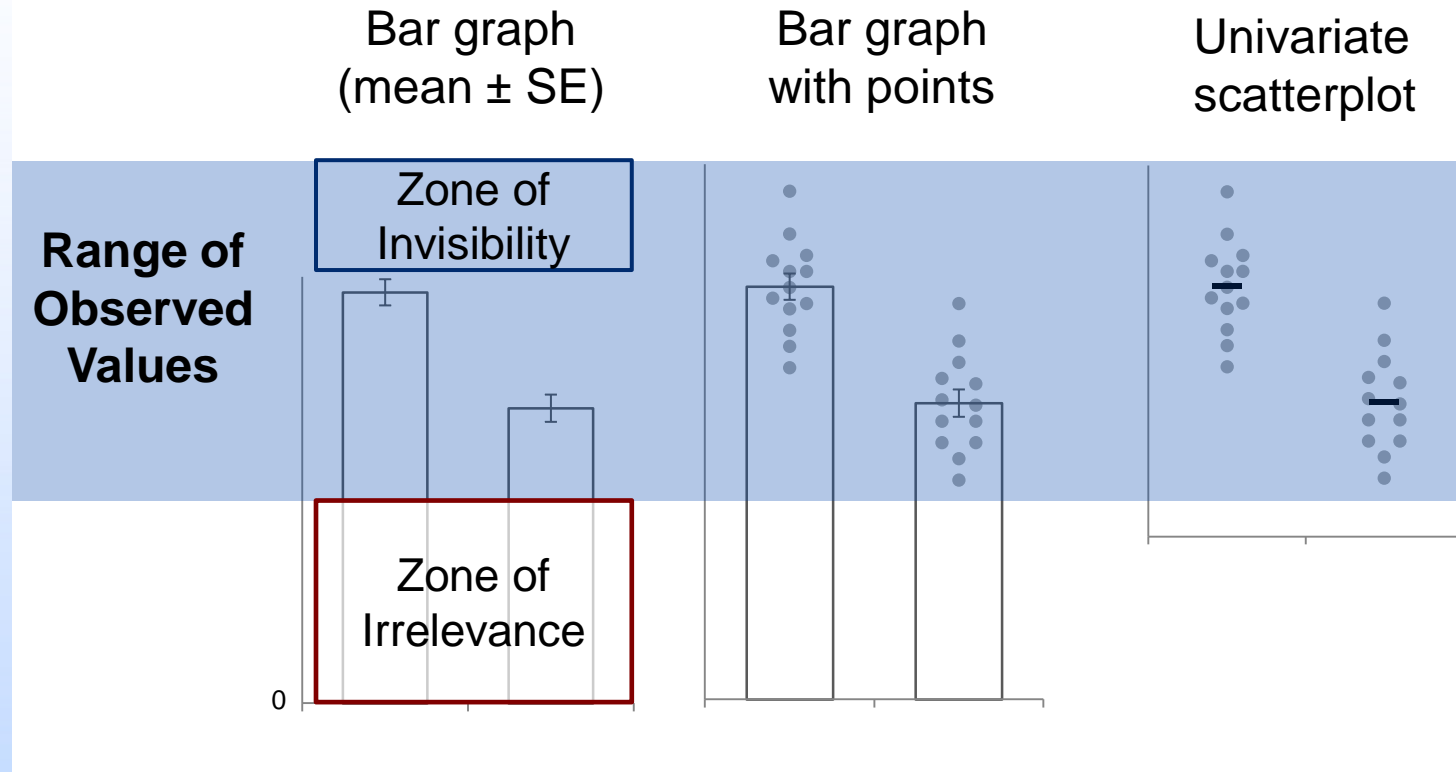
- Almost all papers use bar and line graphs to present continuous data
  - Impossible to critically evaluate the data
- Sample sizes are very small ( $n < 10$  / group)
- 78% of bar graphs show mean SE
- >50% of papers that use non-parametric analysis present data as mean SE or mean SD

# Many different data distributions can lead to the same bar graph...



Test	p value			
T-test: equal var.	0.035	0.074	0.033	0.051
T-test: Unequal var.	0.035	0.076	0.033	0.035
Wilcoxon	0.056	0.10	0.173	0.067

# Why you shouldn't use a bar graph even if your data are normally distributed

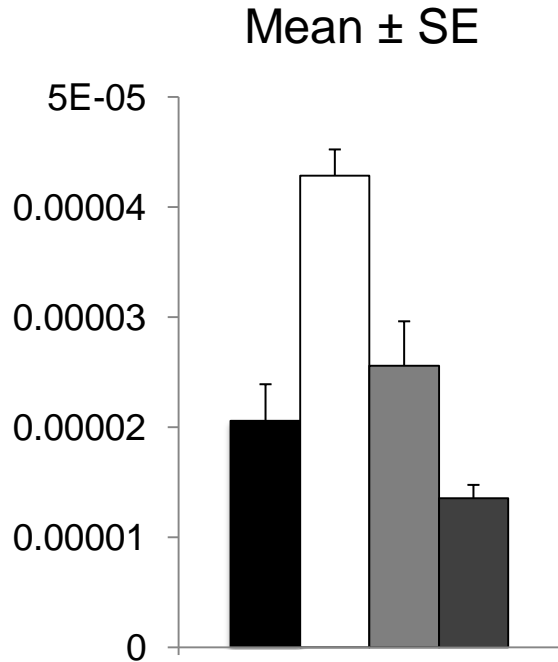


## Bar graphs

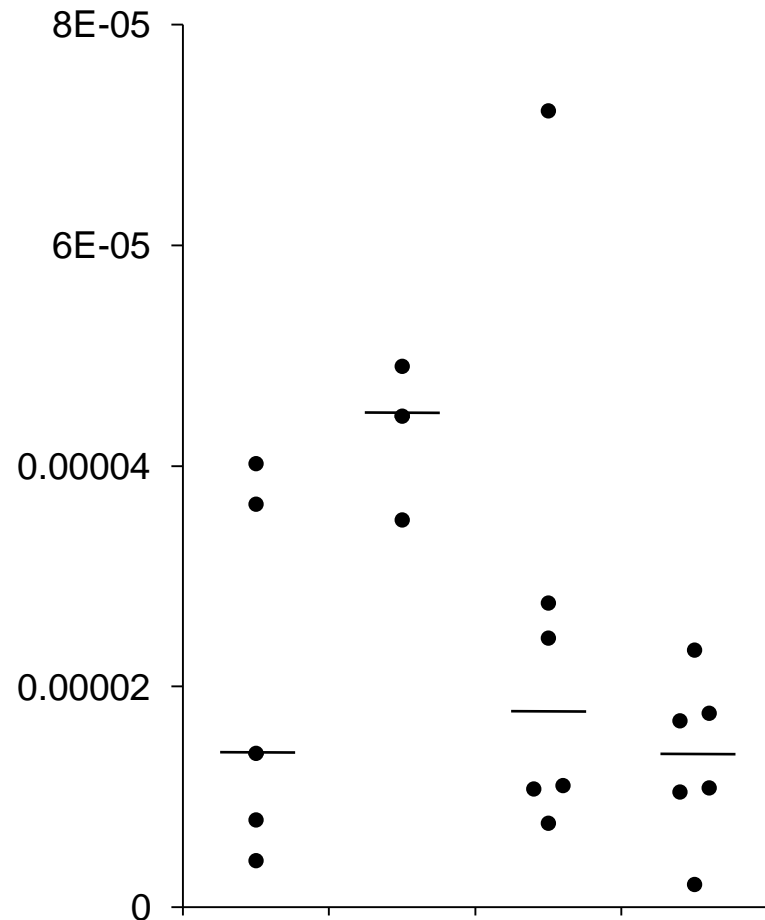
1. Don't allow you to critically evaluate continuous data
2. Arbitrarily assign importance to bar height, rather than focusing on how the difference between means compares to the variability in the data

# Our interpretation depends on what we see

**Reader is a  
passive observer**



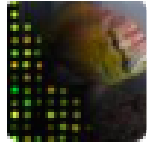
**Reader is an  
active participant**





**Tom Oates** @toates\_19 · Apr 22

Everyone in the world with access to numbers and plot making software should have to read this paper



**Hofmann Lab** @HofmannLabU\_TX · May 2

Finally! Paper in @PLOSbiology makes it plainly obvious why bar graphs hide the data and (often) mislead the reader [bit.ly/1KiLorV](http://bit.ly/1KiLorV)



**Tim Doherty** @TimDoherty\_ · May 12

Just made a bunch of bar plots for my ICCB poster, but then realised I'll probably get shunned after the @PLOSbiology paper



**P. Sean McDonald**

@pseanmc

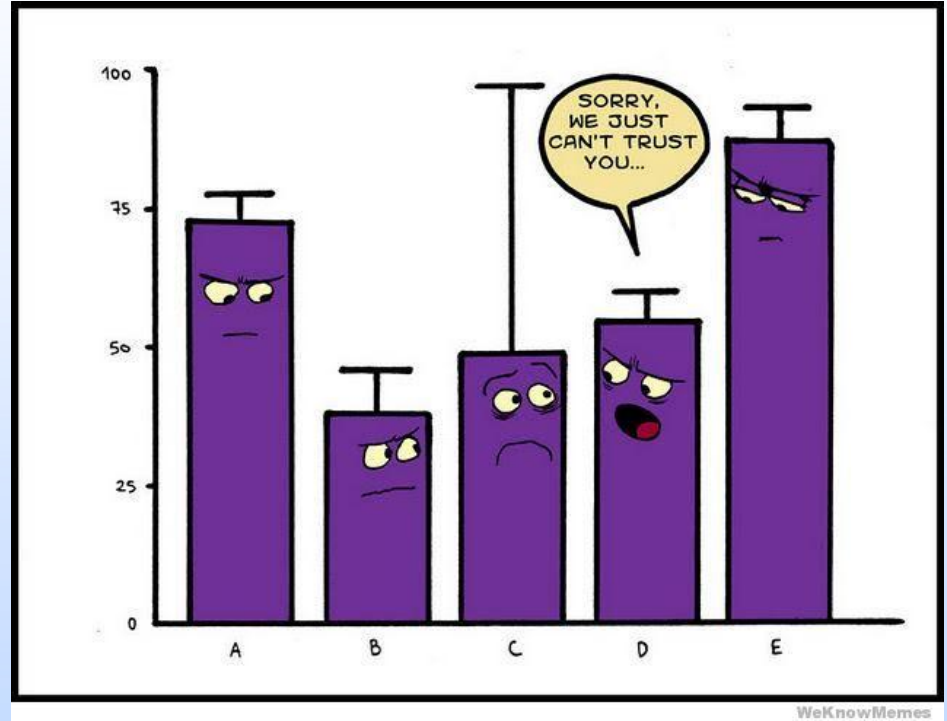
Weissgerber et al., which could be retitled, "Beyond bar graphs: why Sean's graphs suck". [ow.ly/McvM6](http://ow.ly/McvM6)



**Taha Yasseri** @TahaYasseri · Apr 26

Wondering why I'm going to reject your next paper? Reason explained here [bit.ly/1K9vlGo](http://bit.ly/1K9vlGo) awesome @PLOSbiology

# Data presentation is the foundation of our collective scientific knowledge...



Figures are especially important. They often show data for key findings.

Improving data presentation is critical to understanding many diseases

# Why focus on small datasets?

- Common in basic biomedical and biological sciences, translational science
- Influence decisions about what potential treatments advance to clinic trials, further research
- NIH highlighted problems with reproducibility, especially in preclinical research

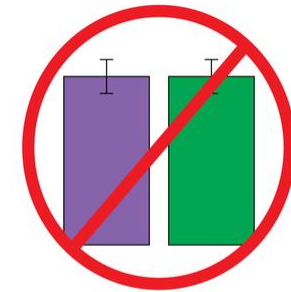


# Must try harder

- Low number of cancer-research studies have been converted into clinical success
- Major factor is the overall poor quality of published preclinical data
- A warning sign should be the “shocking” number of research papers in the field for which the main findings could not be reproduced
- The finding resonates with a growing sense of unease among specialist editors and not just in the field of oncology
- Across the life sciences, handling corrections that have arisen from avoidable errors in manuscripts has become an uncomfortable part of the publishing process

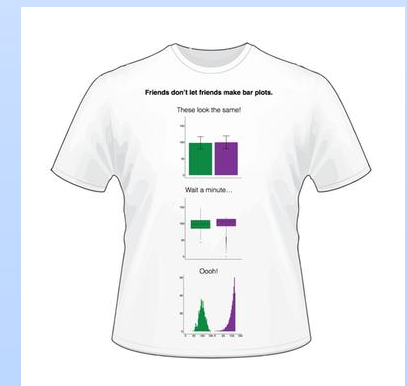
# Effects

- Journal policy changes
  - PLOS Biology
  - Journal of Biological Chemistry
  - Kidney International
  - Journal of Neuroscience Research



#barbarplots

- Editors & reviewers using the paper when requesting improved data visualization
- “Bar bar plots” Kickstarter campaign



# Why this happen?

Statistics are essential,  
but training is not always required for a PhD

Latest articles from *PLOS Biology*

---

April 8, 2016

Gauging the Purported Costs of P...

by Simon Robin Evans

---

April 8, 2016

Reinventing Biostatistics Educatio...

by Natasa M. Milic

<b>135</b> Save	<b>3</b> Citation
<b>14,554</b> View	<b>530</b> Share

PERSPECTIVE

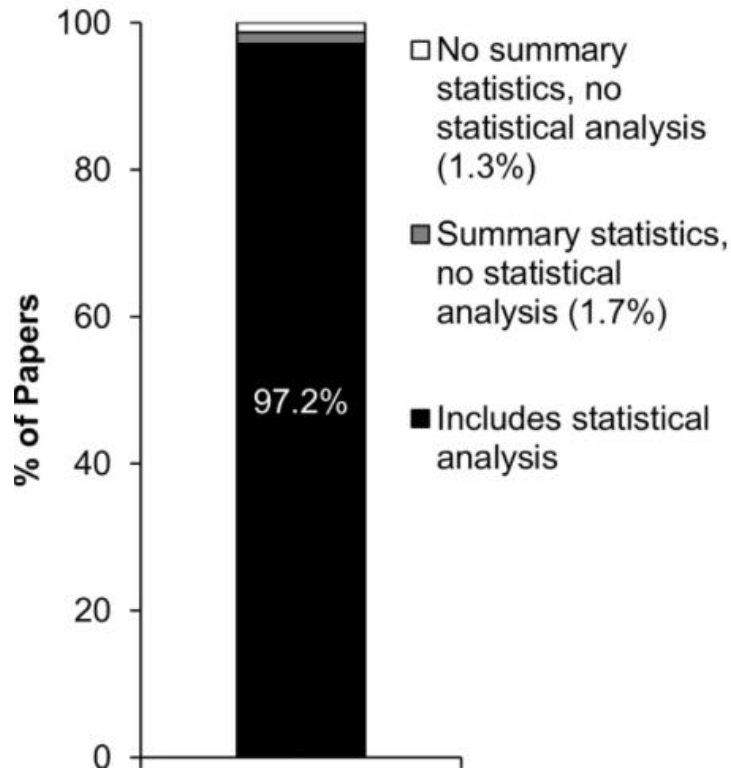
## Reinventing Biostatistics Education for Basic Scientists

Tracey L. Weissgerber<sup>1\*</sup>, Vesna D. Garovic<sup>1</sup>, Jelena S. Milin-Lazovic<sup>2</sup>, Stacey J. Winham<sup>3</sup>, Zoran Obradovic<sup>4</sup>, Jerome P. Trzeciakowski<sup>5</sup>, Natasa M. Milic<sup>1,2</sup>

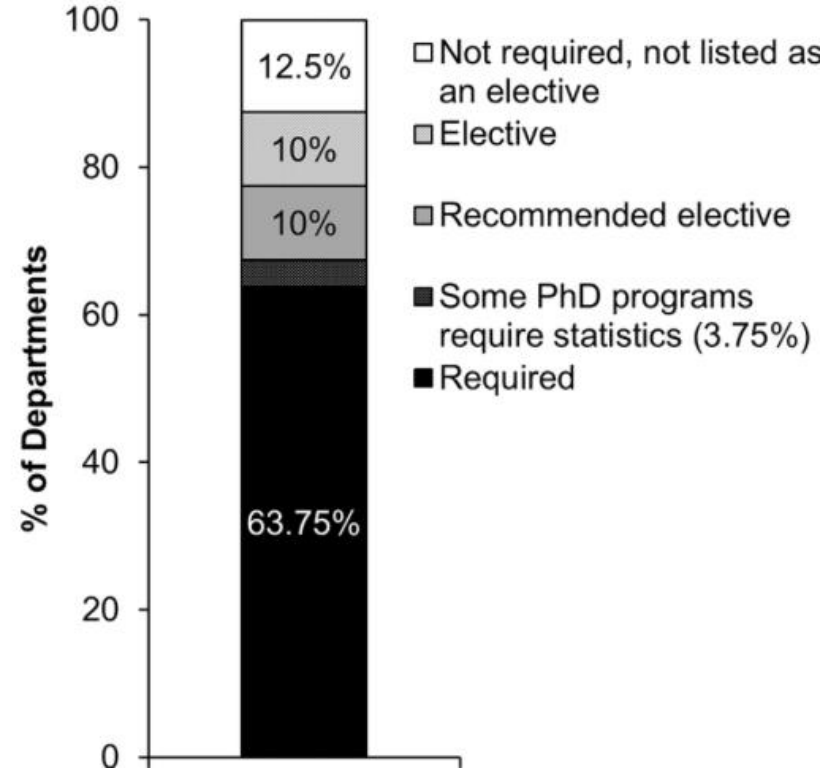
Congratulations,

Out of all 2016 PLOS Biology articles, yours was in the top 50 most downloaded.

**A** Statistics usage among papers published in top physiology journals



**B** Statistics education requirements for PhD programs in top NIH funded physiology departments

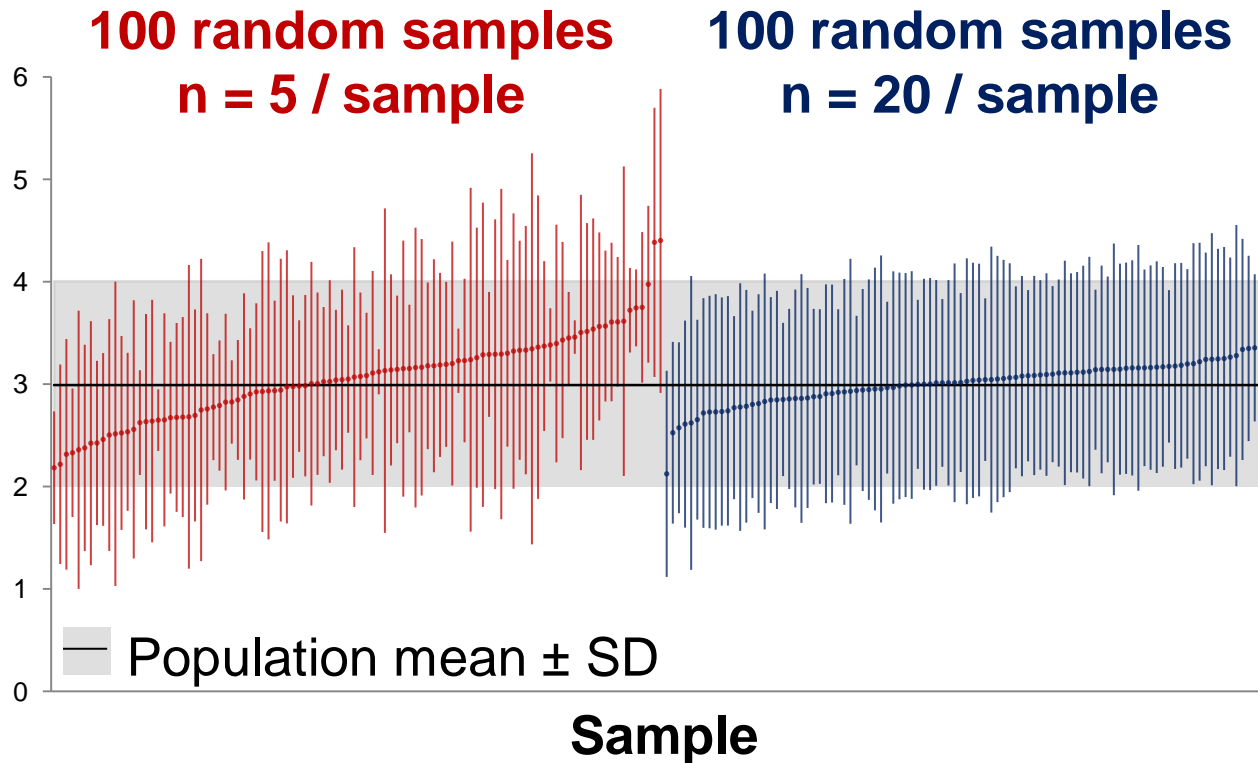


Courses often not targeted towards basic scientists

# Developing new curriculum

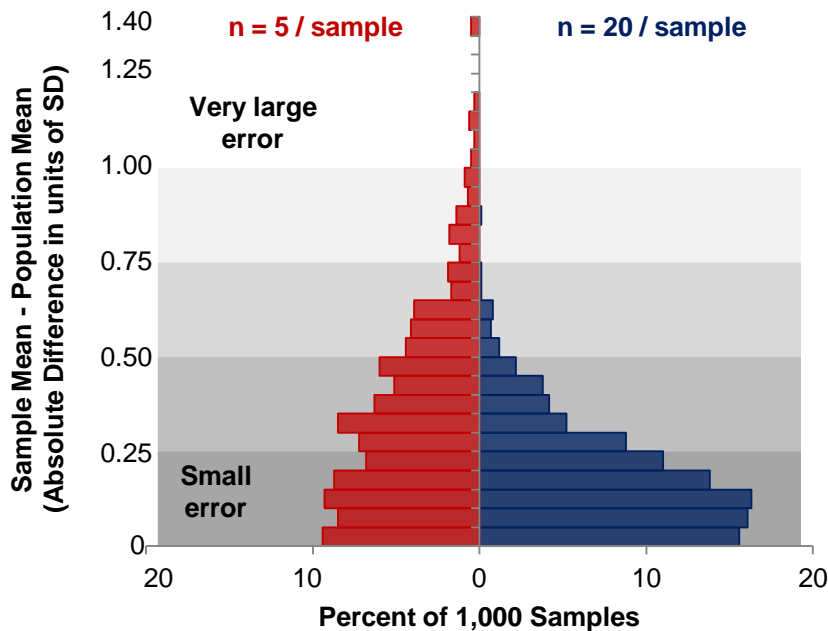
1. Data visualization first, statistics second
2. Target misconceptions & missed skills
3. Visual approach to learning

# Summary statistics are only meaningful when there are enough data to summarize



With small sample sizes, means and SDs can be very inaccurate

### Error in Sample Means

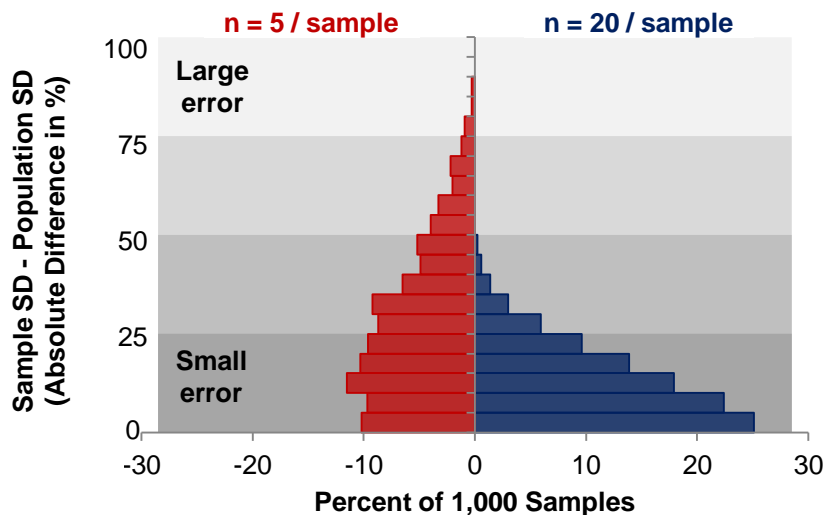


Error in Sample Mean	n = 5 / sample	n = 20 / sample
>1 SD	2.2%	0.0%
0.75 to 1 SD	6.0%	0.1%
0.5 to 0.75 SD	16.0%	2.9%
0.25 to 0.5 SD	33.1%	24.2%
< 0.25 SD	42.7%	72.8%

Sample mean is >1 SD away from population mean

Sample mean is within 0.25 SD of population mean

### Error in Sample SD



Error in Sample SD	n = 5 / sample	n = 20 / sample
75 to 100%	1.5%	0.0%
50 to 75%	12.7%	0.0%
25 to 50%	34.5%	11.1%
< 25%	51.3%	88.9%

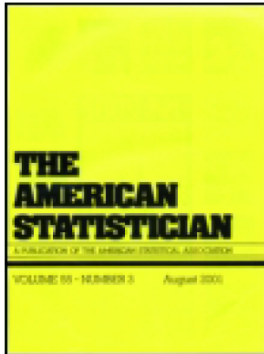
Sample SD is 75 to 100% different from population SD

Sample SD is within 25% of population SD



# Reproducibility in focus

- No one denies that irreproducibility of scientific research is a serious problem
- Statistical community concern
- Basic and Applied Social Psychology: banded p-values (Trafimow and Marks, 2015)
- “Reproducibility crisis” (Peng, 2015)
- This concern was brought to the attention of the ASA Board



## The American Statistician

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: <http://www.tandfonline.com/loi/utas20>

# The ASA's statement on p-values: context, process, and purpose

Ronald L. Wasserstein & Nicole A. Lazar

To cite this article: Ronald L. Wasserstein & Nicole A. Lazar (2016): The ASA's statement on p-values: context, process, and purpose, The American Statistician, DOI: [10.1080/00031305.2016.1154108](https://doi.org/10.1080/00031305.2016.1154108)

To link to this article: <http://dx.doi.org/10.1080/00031305.2016.1154108>

## Let's be clear

- Nothing in the ASA statement is new.
- Statisticians have been sounding the alarm about these matters for decades, to little avail.
- We hoped that a statement from the world's largest professional association of statisticians would open a fresh discussion and draw renewed and vigorous attention to changing the practice of science with regards to the use of statistical inference.

# What is a p-value?

- Informally, a p-value is the probability under a specified statistical model that a statistical summary of the data (for example, the sample mean difference between two compared groups) would be equal to or more extreme than its observed value.

# Principles

- *P-values can indicate how incompatible the data are with a specified statistical model*
- *P-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.*
- *Scientific conclusions and business or policy decisions should not be based only on whether a p-value passes a specific threshold.*

# Principles continued

- *Proper inference requires full reporting and transparency*
- *A p-value, or statistical significance, does not measure the size of an effect or the importance of a result*
- *By itself, a p-value does not provide a good measure of evidence regarding a model or hypothesis.*

# Other approaches

- Methods that emphasize estimation over testing, such as confidence, credibility, or prediction intervals; Bayesian methods; alternative measures of evidence, such as likelihood ratios or Bayes Factors; and other approaches such as decision-theoretic modeling and false discovery rates

# Editor's Perspective

## The End of Journals

Harlan M. Krumholz, MD, SM

### Too Static

The journal publication is currently a static product, presented as a singular contribution rather than as a living document. It can be corrected or retracted, but it is not interactive and has no capacity for iterative change spurred by input from the larger audience. Many scientific projects might be better presented as an interactive website with the opportunity for the community to probe the findings and provide feedback. Creative visualizations of data are often best presented in ways that allow images to be rotated and manipulated for better understanding. In the future, novel strategies for conveying knowledge and engaging readers will probably emerge, leaving behind the static presentation of results that offers limited options for interactive understanding.





COMMUNITY PAGE

## From Static to Interactive: Transforming Data Visualization to Improve Transparency

Tracey L. Weissgerber<sup>1\*</sup>, Vesna D. Garovic<sup>1‡</sup>, Marko Savic<sup>2</sup>, Stacey J. Winham<sup>3</sup>, Natasa M. Milić<sup>1,2‡</sup>

### Objective:

Create tools needed to transform scientific publications from

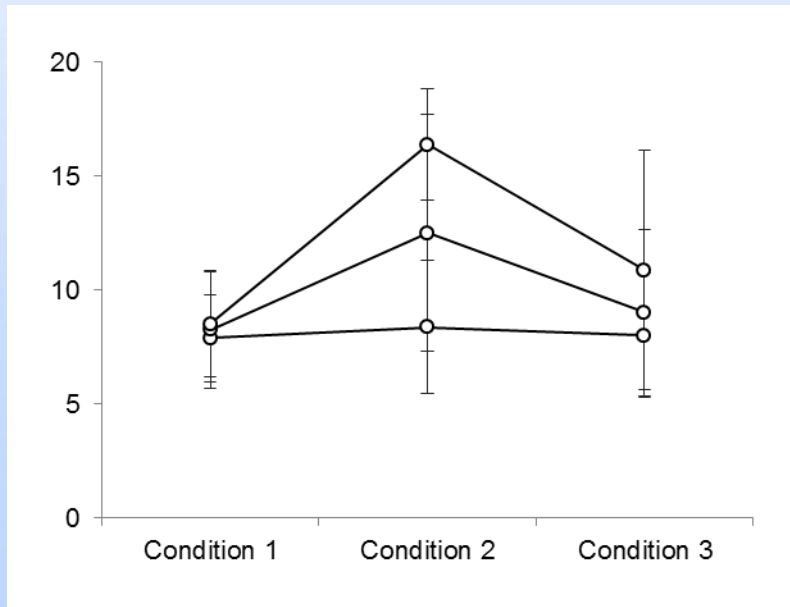
*static reports into interactive datasets*

# Effective figures should:

1. Immediately convey information about the study design
2. Illustrate important findings
3. Allow the reader to critically evaluate the data

# Example: The line graph

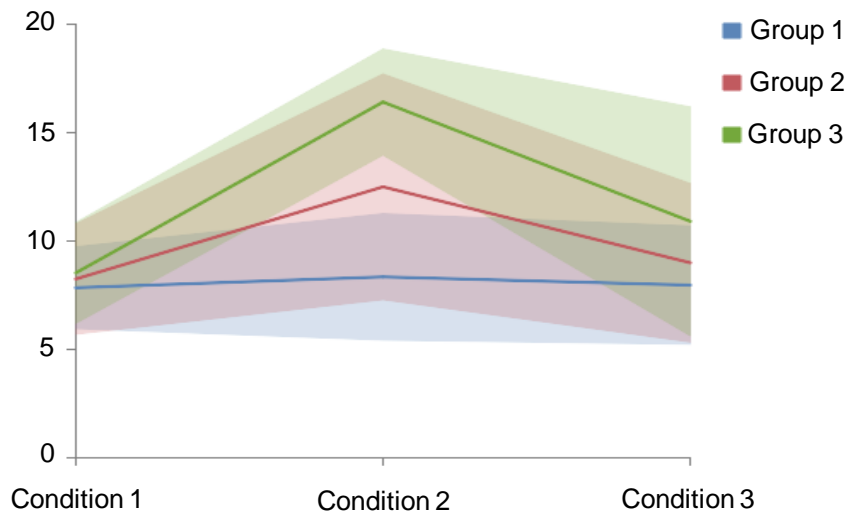
- Scientists want to know 2 things:
  1. How much do different groups overlap?
  2. Do all individuals in a group follow the same response pattern?



Difficult or impossible to determine with a static line graph

# Solution: Interactive Line Graph

<http://statistika.mfub.bg.ac.rs/interactive-graph/>



## Interactive Line Graph

Please cite this tool as [Video instructions](#)

**Link to video  
tutorial on site**

The interactive line graph allows you to do four things:

1. Examine different summary statistics

2. Display lines for some or all individuals in each group

3. View a subset of groups, conditions or time points

4. View change scores for any two conditions or time points

# Two Options for Easy Data Entry

## 1. Manual data entry

Data

Define your dataset

<b>Number of groups:</b> <input type="text" value="2"/>	<b>Number of conditions:</b> <input type="text" value="3"/>
--	--

Up to 6 groups      Up to 10 conditions    [Need more?](#)

<b>Group 1 label:</b> <input type="text" value="Placebo"/>	<b>Condition 1 label:</b> <input type="text" value="0 weeks"/>
<b>group 1 size:</b> <input type="text" value="10"/>	<b>Condition 2 label:</b> <input type="text" value="1 week"/>
<b>Group 2 label:</b> <input type="text" value="Drug"/>	<b>Condition 3 label:</b> <input type="text" value="2 weeks"/>
<b>group 2 size:</b> <input type="text" value="10"/>	

Save and proceed to enter data     Don't save, just download template CSV file



Download template csv file

## 2. Upload csv file

Upload workset

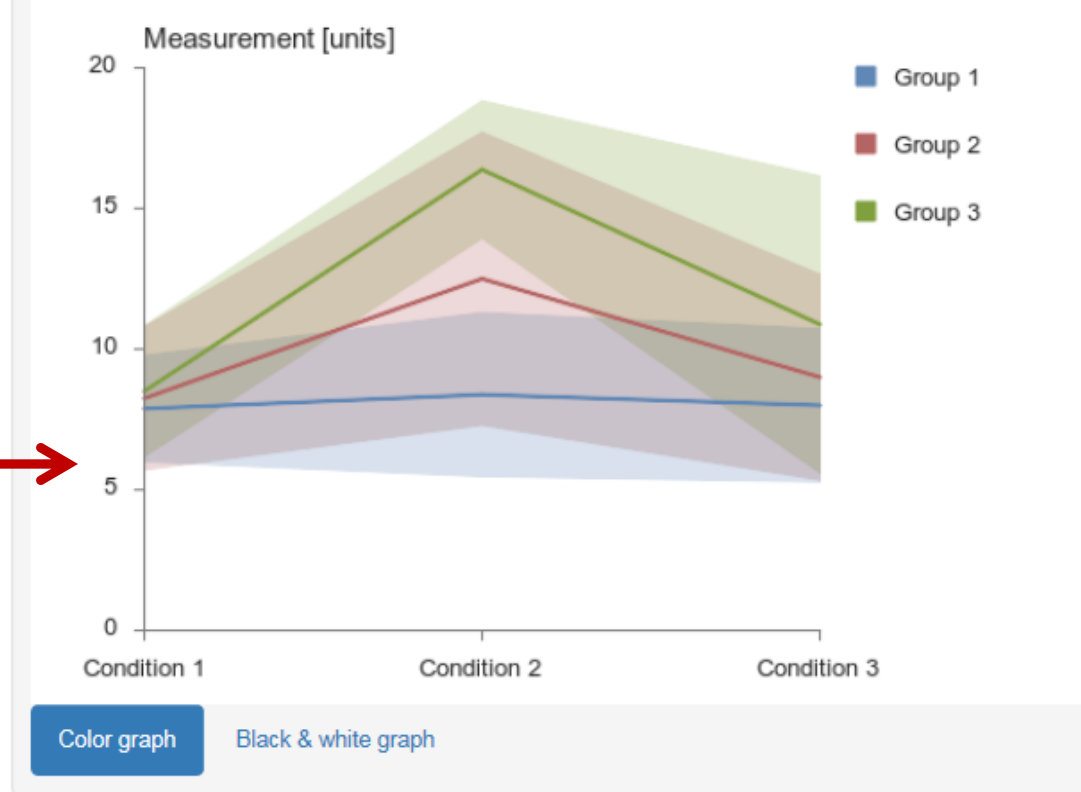
Upload your workset

**XML or CSV file:**

# View different summary statistics

Semi-transparent shading makes it easy to assess overlap between groups

Select any option; the graph will automatically update



Graph data

Present data as

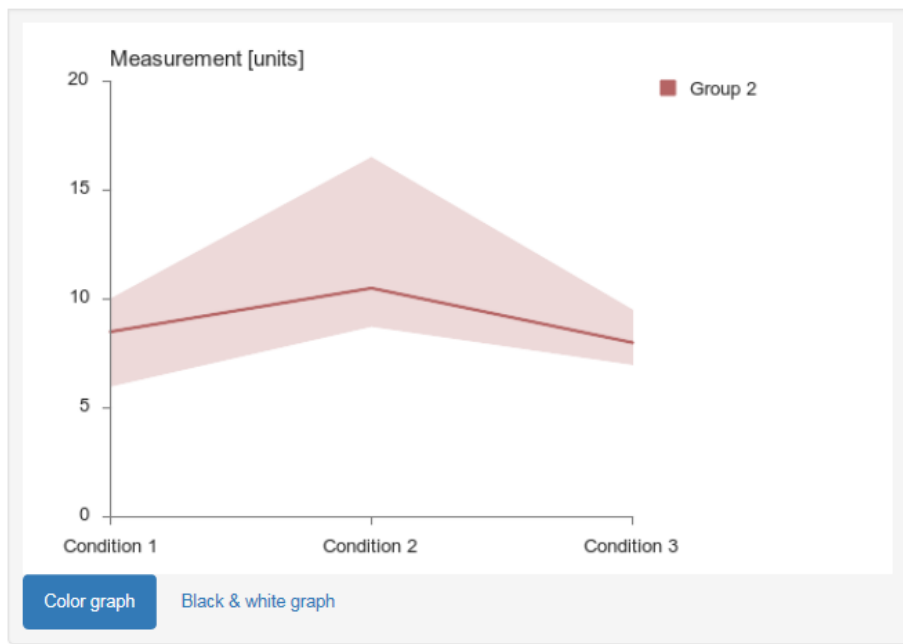
- mean
- mean (SD)
- mean (SE)
- mean (95% CI)
- med
- med (IQR)
- med (range)

Show conditions  as continuous time points

- Condition 1
- Condition 2
- Condition 3

# Focus on specific groups, conditions or time points

Uncheck the box next to any group or condition to delete it from the graph



Graph data

Present data as

- mean
- mean (SD)
- mean (SE)
- mean (95% CI)
- med
- med (IQR)
- med (range)

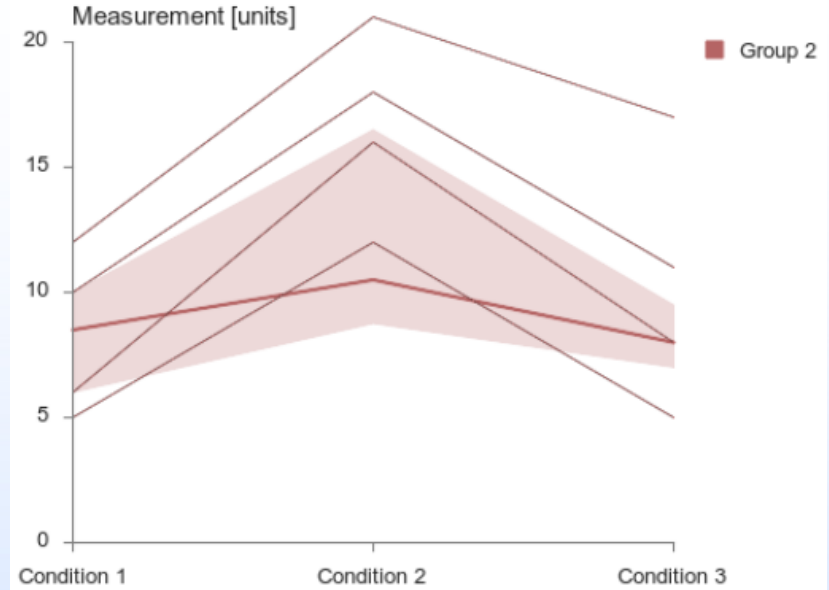
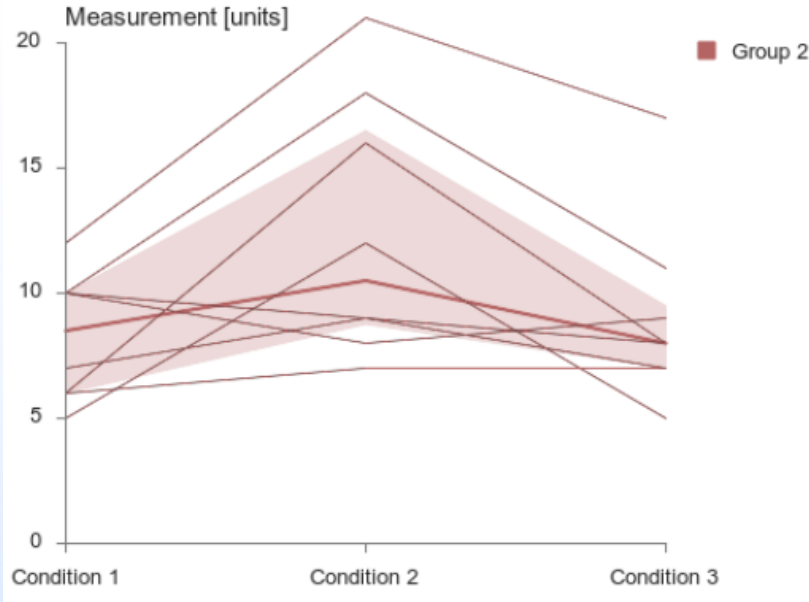
Show conditions  as continuous time points

- Condition 1
- Condition 2
- Condition 3

Group 1     Group 2     Group 3



# View lines for any individual



Check boxes to view individual lines



Group 2

Select All/None

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.

Group 2

Select All/None

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.

# View change scores for any 2 conditions / time points

Difference plot tab →



Change summary line →

Change conditions →

Show / hide groups →

Graph data

Central tendency measure:

Mean

Median

Select conditions:

Condition 1

Condition 2

Show groups:

Group 1

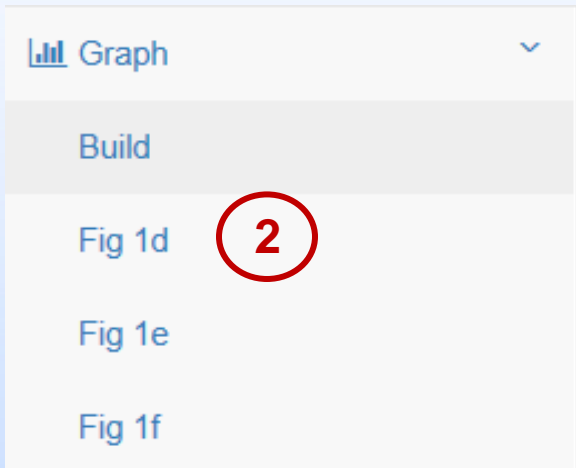
Group 2

Group 3

# Features for Publications

1

Save this graph as:



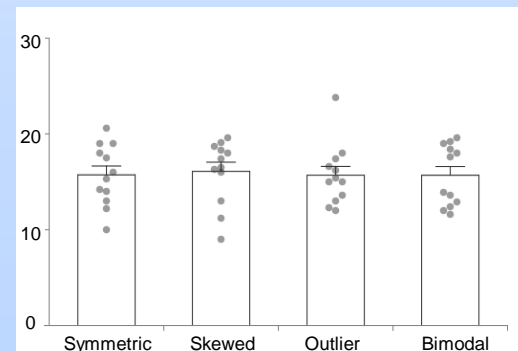
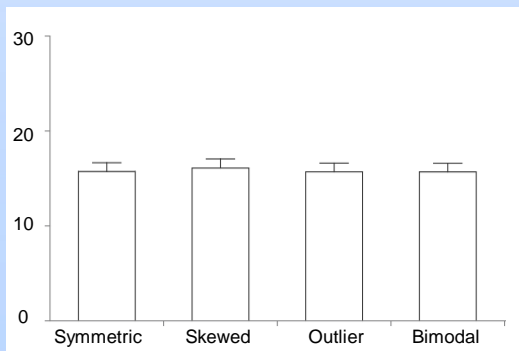
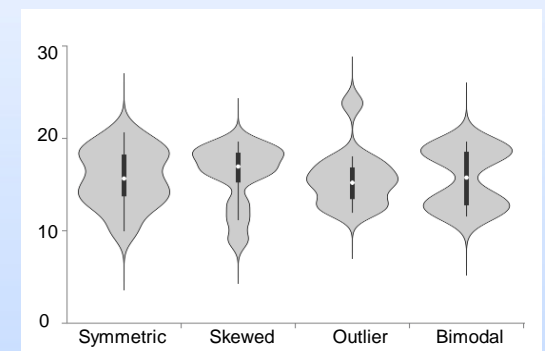
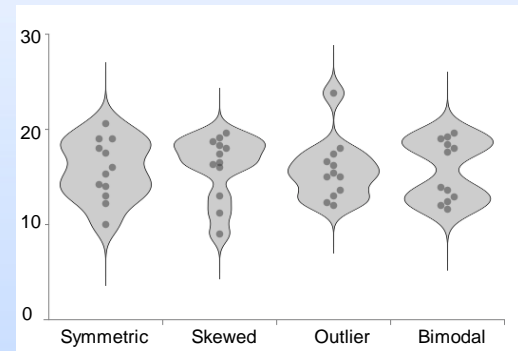
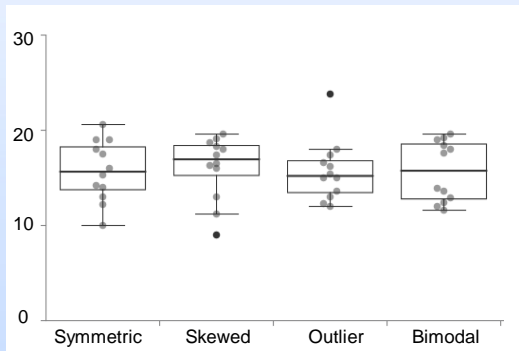
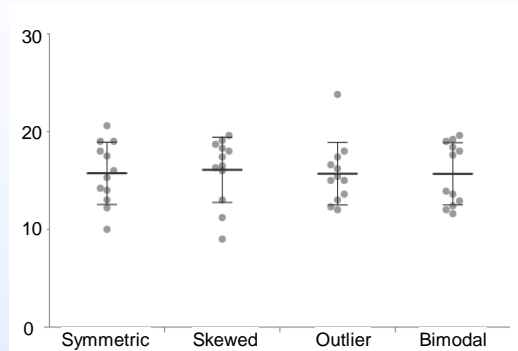
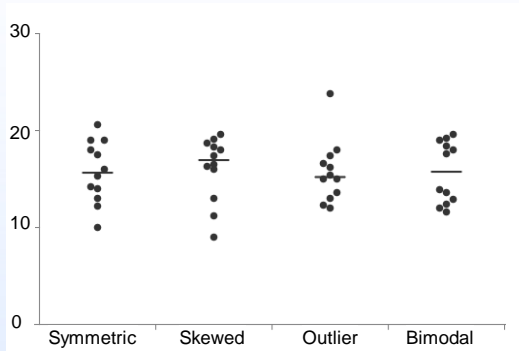
 Download workset

3

1. Save any static graph that you want to include in your paper
2. Saved graphs appear in the “Build” menu. Readers can explore the entire dataset or click on a saved graph to explore a specific figure from your paper.
3. You can download:
  - csv datafile
  - Tiff files for any saved static graphs
  - Xml file containing your interactive graphic to include in the supplement of your paper

# Interactive Dotplot (in press)

<http://statistika.mfub.bg.ac.rs/interactive-dotplot/>



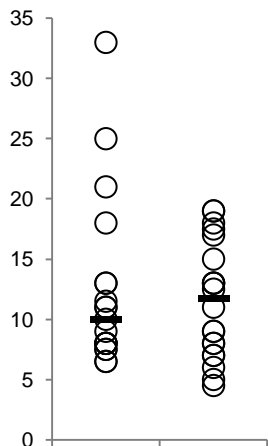
**Warning:** The bar graph is included as a teaching tool and is not intended for use in scientific publications. See relevant publications [1] [2] and initiatives [1]

# Making effective dotplots

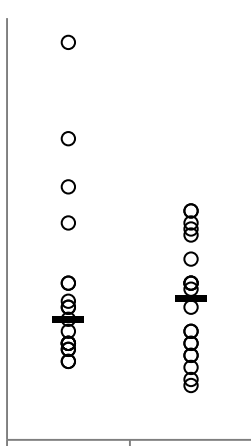
- Dotplots are useful for comparing values of a continuous variable across study groups (cross-sectional study or experimental study with independent groups)
- Dotplots are the best choice for small datasets. The summary statistics shown in box plots and violin plots are only meaningful if you have enough data to summarize.

**Step 1:**  
Make all  
data points  
visible

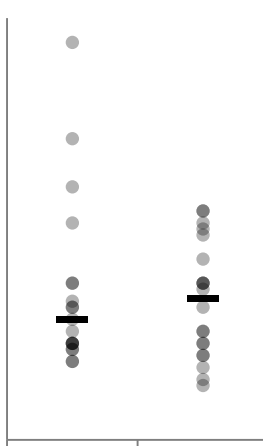
**A:** Ineffective graph



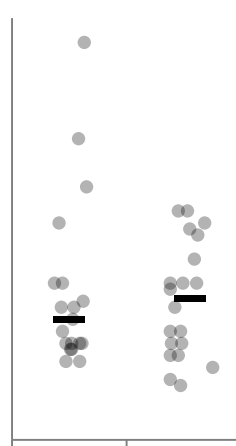
**B:** Decrease point size



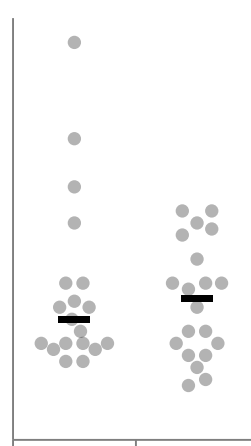
**C:** Semi-transparent points



**D:** Random jitter

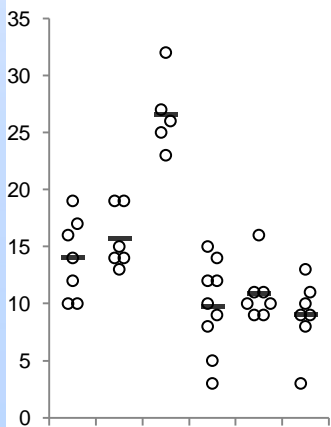


**E:** Symmetric jitter

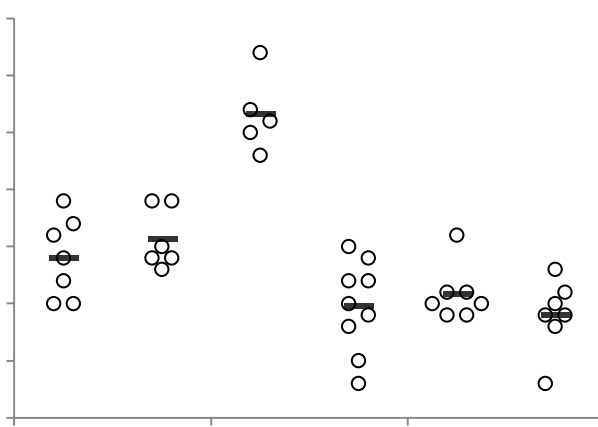


**Step 2:**  
Emphasize  
summary  
statistics

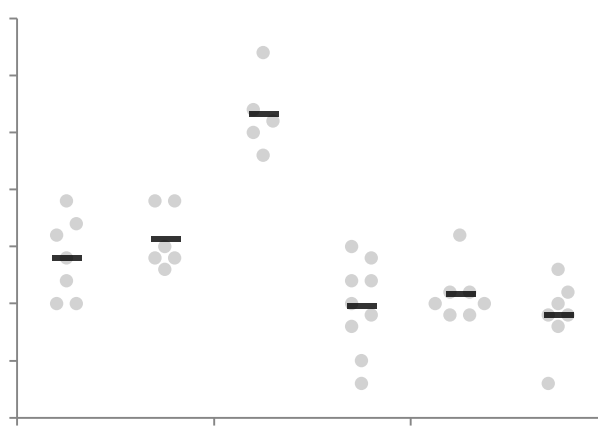
**A:** Ineffective graphic



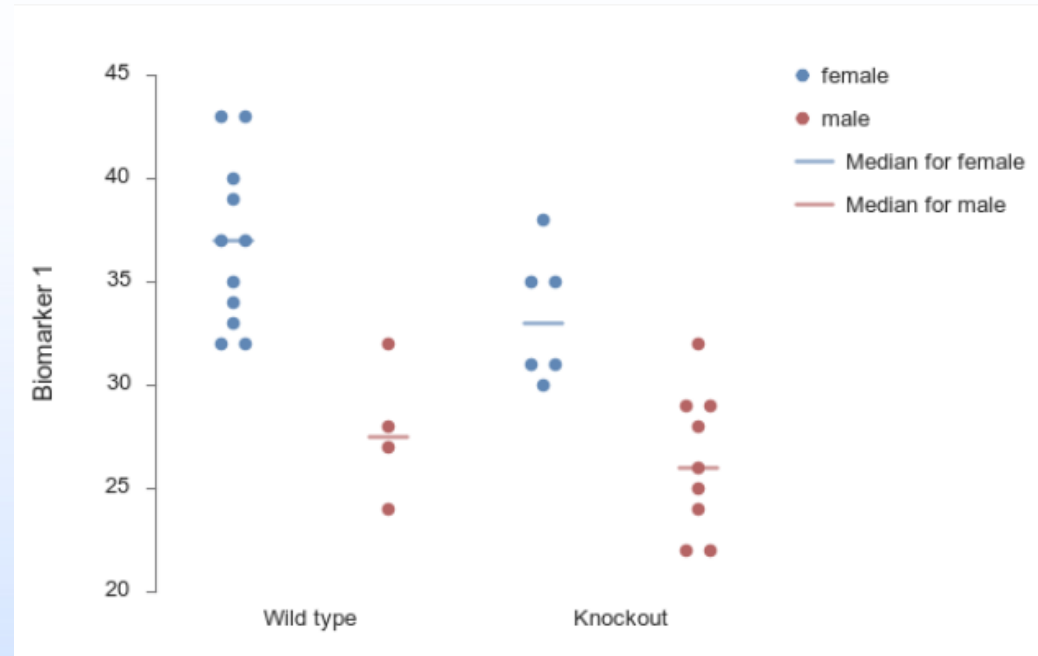
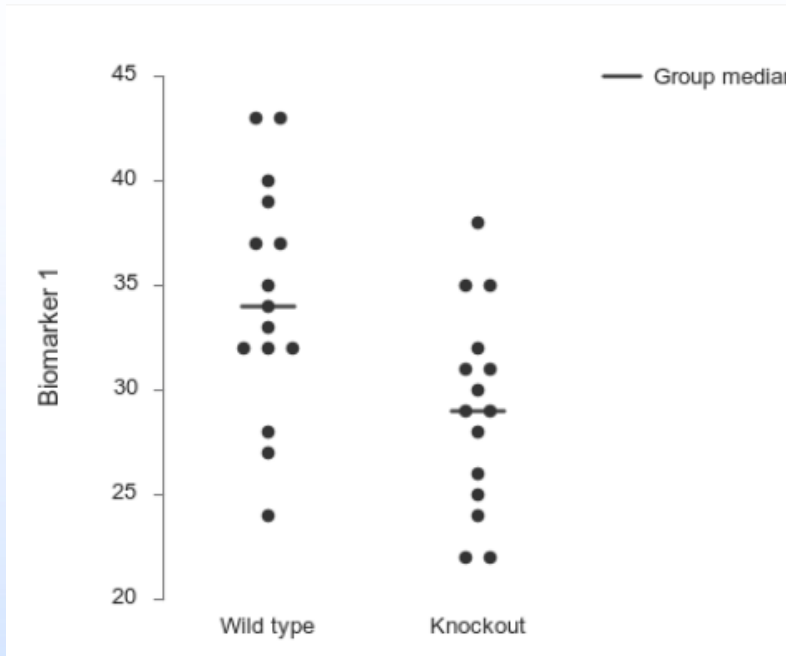
**B:** Increase width



**C:** Emphasize summary statistics



# Examine subgroups (i.e. male vs. female)



Subgroups:  color code

Show subgroups side by side

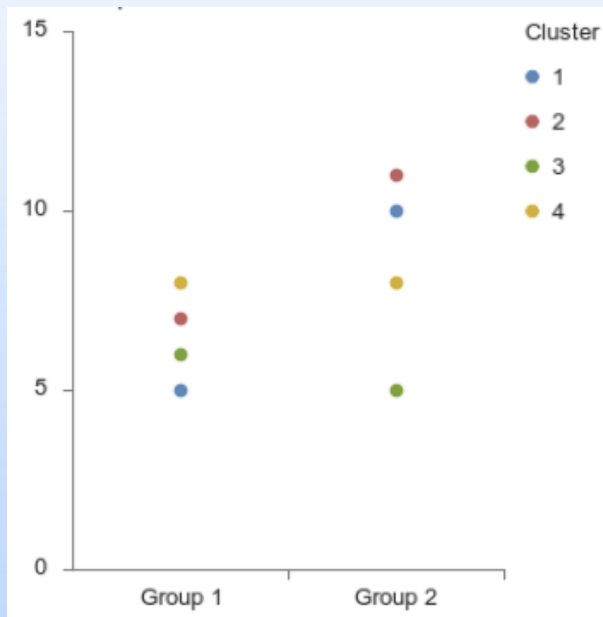
Summary lines: Median

female

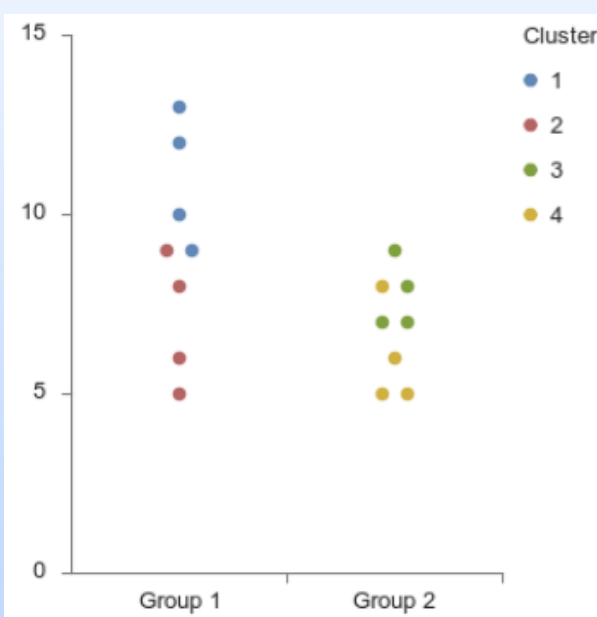
male

# Show clusters of non-independent data (replicates, mice from the same litter)

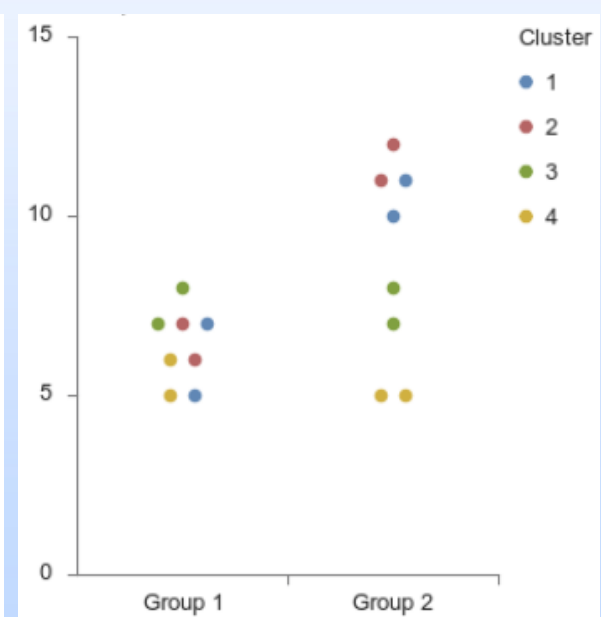
## Between-Group Clusters



## Within-Group Clusters

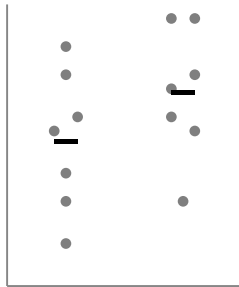


## Between & Within-Group Clusters

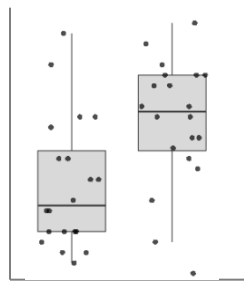


# Graphics for:

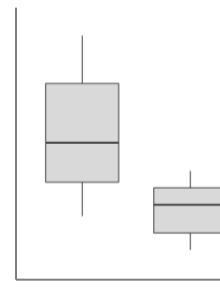
- Cross sectional studies
- Experimental studies with independent groups



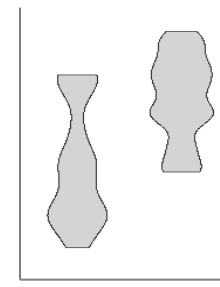
**Dotplot**



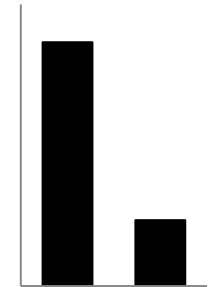
**Boxplot with points**



**Boxplot**



**Violin plot (with or without points)**



**Bar graph**

Outcome variable	Continuous	Continuous	Continuous	Continuous	Counts & proportions
Sample size	Small	Medium	Large	Medium to Large	Any
Data distribution	Any	Any	Do not use for bimodal data	Any	N/A

Free violin plot tool: <https://interactive-graphics.shinyapps.io/violin/>



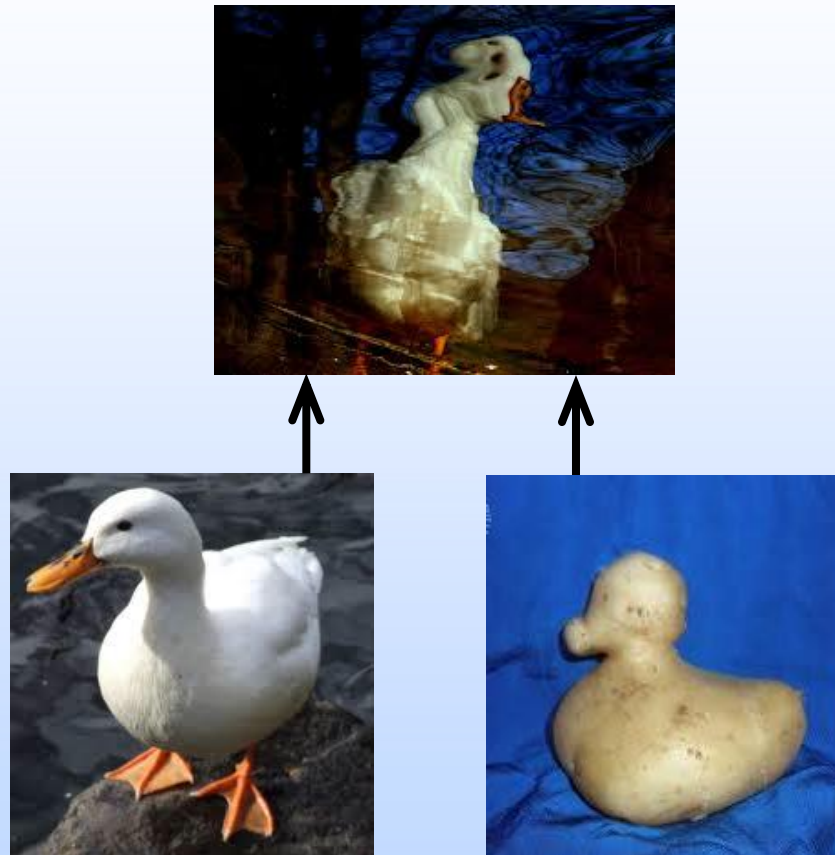
# Two Types of Solutions

- Author level
  - Small interactive graphic program file that authors can upload in data supplement of a published paper
- Journal level
  - Interactive graphic in online manuscript
  - User customizes data presentation before printing
- Ensure compatibility

# Changing data presentation is critical to promote transparency



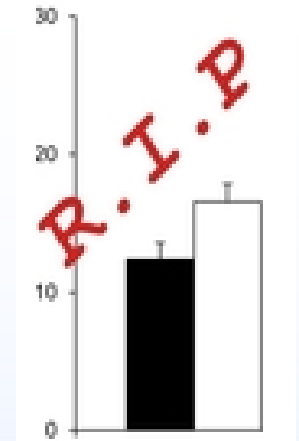
All data presentation methods are a reflection of reality...



Select methods that minimize distortion

# What can you do?

- Banish bar graphs from your papers and talks
- Reviewers & editors: Request figures that show data distributions
- Talk to editors about improving data presentation in their journals
- Work with statistics instructors to organize data presentation training for trainees, junior investigators and senior researchers





## Announcement: Towards greater reproducibility for life-sciences research in *Nature*

May 2017

This week we go further. Alongside every life-sciences manuscript, we will publish a [new reporting-summary document](#).

This is another step in encouraging transparency, in ensuring that papers contain sufficient methodological detail, and in improving statistics reviewing and reporting.

New policy includes a variety of steps to improve reporting & transparency, including discouraging the use of bar graphs and encouraging authors to show the data distribution.

# Efforts are under way

- Initiatives by journals
- Pre-publication efforts

# Conclusion

- Good statistical practice, as an essential component of good scientific practice, emphasizes principles of good study design and conduct, a variety of numerical and graphical summaries of data, understanding of the phenomenon under study, interpretation of results in context, complete reporting and proper logical and quantitative understanding of what data summaries mean.
- No single index should substitute for scientific reasoning.

# Acknowledgements

## **Mayo Clinic**

Vesna Garovic

Tracey Weissgerber

Stacey Winham

**NIH BIRCWH K12HD065987**

**NCATS UL1 TR000135**

**NICHD P-50 AG44170**

## **University of Belgrade**

Marko Savic

Jelena Milin-Lazovic

## **Temple University**

Zoran Obradovic