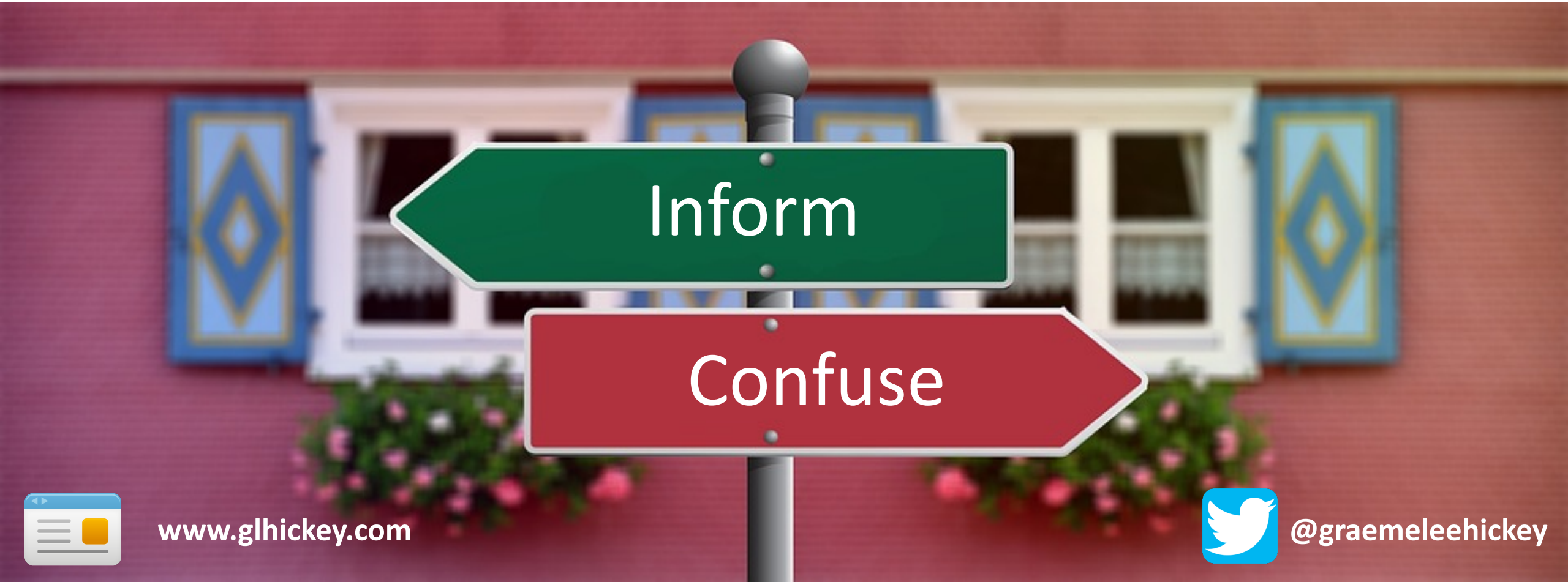
To inform or confuse with tables and figures: the EJCTS experience

Graeme L. Hickey
University of Liverpool



www.glhickey.com

@graemeleehickey

Conflicts of interest

- None
- Assistant Editor (Statistical Consultant) for EJCTS and ICVTS

Summarizing data

- Very small number of statistics – report in-line
  - E.g. “The in-hospital mortality was 10% (*n* = 20)”

- Many unrelated statistics (e.g. different patient characteristics) or displaying fine-level detail – report in tabular format

- Many related statistics (e.g. biomarker values over time) or data to complex for modelling – report in graphical format

Figures as the natural presentation tool

Flowcharts

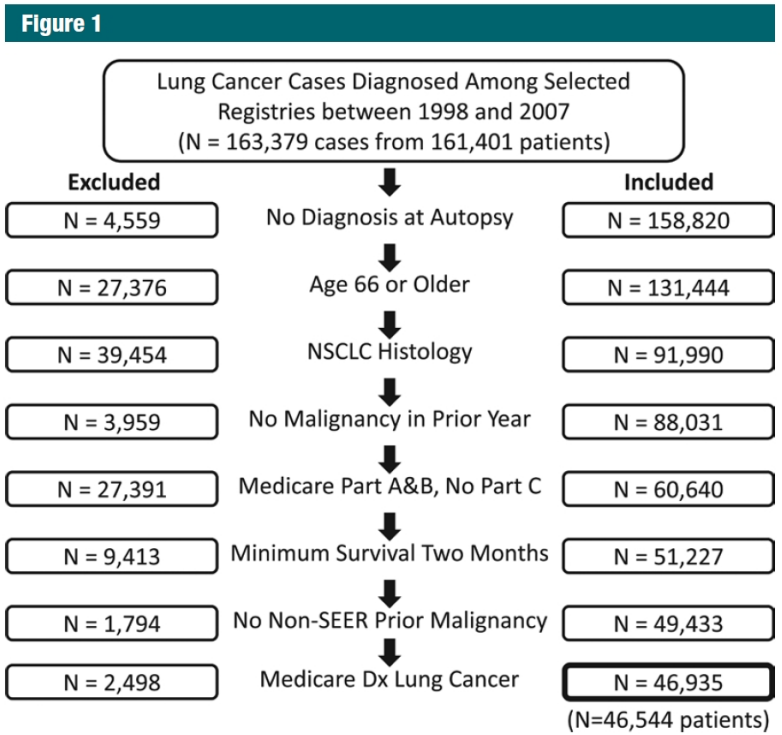
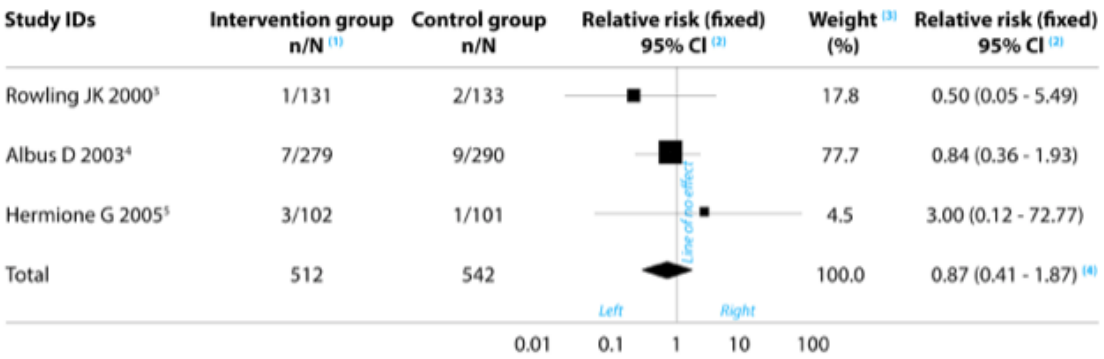


Figure 1: Flow diagram of study inclusion and exclusion criteria. *Dx* = diagnosis.

Source: Benchimol et al. *PLoS Med* 2015; 12(10): e1001885.

Forest plots



Test for heterogeneity Chi-square = 0.79, df = 2, $p = 0.67$, $I^2 = 0.0\%$ (5)
Test for overall effect $z = 0.35$, $p = 0.7$ (6)

(1) N = total number in group, n = number in group with the outcome.
(2) Outcome of interest in picture and in number. Fixed effect model used for meta-analysis.
(3) Influence of studies on overall meta-analysis.
(4) Overall effect.
(5) Heterogeneity (I^2) = 0%. So, we use fixed effect model.
(6) *p* value indicating level of statistical significance

Source: http://uk.cochrane.org/news/how-read-forest-plot

Tables as the natural presentation tool

Summarizing + comparing data of different types

Table 1: Patient and operative characteristics data by CPB technique with statistical comparison

	Overall		On-pump		Off-pump		Δ (%)	P
Total number	n = 3402		n = 1173		n = 2229			
Logistic EuroSCORE (%)	2.4 ± 2.5		2.4 ± 2.8		2.3 ± 2.3		1.8	0.965
Age (years)	61.7 ± 10.6		61.1 ± 10.3		61.9 ± 10.7		−8.1	0.026
BMI (kg/m^2)	28.5 ± 4.6		28.7 ± 4.7		28.4 ± 4.5		6.1	0.090
	N	%	N	%	N	%		
Female	880	25.9	325	27.7	555	24.9	6.4	0.083
Preoperative AF	69	2.0	28	2.4	41	1.8	3.8	0.343
Urgent	733	21.5	271	23.1	462	20.7	5.7	0.119
NYHA III/IV	645	19.0	225	19.2	420	18.8	0.9	0.846
History of neurological dysfunction	53	1.6	25	2.1	28	1.3	6.8	0.070
Diabetes (insulin- or diet-controlled)	600	17.6	207	17.6	393	17.6	0.0	>0.999
History of hypertension	2269	66.7	764	65.1	1505	67.5	−5.1	0.172
Recent MI	480	14.1	177	15.1	303	13.6	4.3	0.255
Creatinine >200 μmol//l	33	1.0	11	0.9	22	1.0	−0.5	>0.999
History of pulmonary disease	361	10.6	115	9.8	246	11.0	−4.0	0.293
Extracardiac arteriopathy	226	6.6	89	7.6	137	6.1	5.7	0.126
Previous PCI	815	24.0	299	25.5	516	23.1	5.5	0.139
Left ventricular function								
Good (LVEF >50%)	3004	88.3	1011	86.2	1993	89.4	−9.9	0.018
Fair (LVEF 30–50%)	355	10.4	146	12.4	209	9.4	9.9	
Poor (LVEF <30%)	43	1.3	16	1.4	27	1.2	1.4	
Critical preoperative state	34	1.0	15	1.3	19	0.9	4.2	0.314
Preoperative IV nitrates or heparin for treatment of unstable angina	41	1.2	15	1.3	26	1.2	1.0	0.904

BMI: body mass index; MI: myocardial infarction; PCI: percutaneous coronary intervention; LVEF: left ventricular ejection fraction; CVD: coronary vessel disease; IV: intravenous; CABG: coronary artery bypass graft; AF: atrial fibrillation; NYHA: New York Heart Association.
Statistics reported as mean ± standard deviation for continuous variables, and number (%) for categorical/binary variables.

Δ is the standardized difference: $100(\bar{x}_{on} - \bar{x}_{off})/\sqrt{(s^2_{on} + s^2_{off})/2}$, where $\bar{x}_{off}$ and $\bar{x}_{on}$ denote the sample means for the off- and on-pump groups, respectively, and s^2_{off} and s^2_{on} the respective sample variances.
P is the *P*-value: χ^2 test for all categorical variables (some with Yates' continuity correction as appropriate); independent samples *t*-test for age and BMI; the Mann–Whitney *U*-test for logistic EuroSCORE.

Source: Hickey GL et al. *EJCTS*. 2015; 49: 1441–1449.

Summarizing the results of a regression model when the exact coefficients are required

Table 6: Final risk factors by multivariate regression for the model

Risk factor	Coefficient	Standard error	z	P ≥ |z|	[95% confidence interval]
NYHA					
II	0.1070545	0.1463849	0.73	0.465	[−0.1798547, 0.3939637]
III	0.2958358	0.141466	2.09	0.037	[0.0185674, 0.5731042]
IV	0.5597929	0.1697565	3.30	0.001	[0.2270763, 0.8925095]
CCS4	0.2226147	0.1462888	1.52	0.128	[−0.0641061, 0.5093356]
IDDM	0.3542749	0.145863	2.43	0.015	[0.0683887, 0.6401611]
Age	0.0285181	0.0065954	4.32	0.000	[0.0155914, 0.0414448]
Female	0.2196434	0.0953505	2.30	0.021	[0.0327599, 0.4065269]
ECA	0.5360268	0.1106046	4.85	0.000	[0.3192458, 0.7528079]
CPD	0.1886564	0.1232126	1.53	0.126	[−0.0528358, 0.4301486]
N/M mob	0.2407181	0.1729494	1.39	0.164	[−0.0982564, 0.5796927]
Redo	01.118599	0.1226272	9.12	0.000	[0.8782539, 1.3589440]
Renal dysfunction					
On dialysis	0.6421508	0.3083468	2.08	0.037	[0.0378021, 1.2464990]
CC ≤ 50	0.8592256	0.1446758	5.94	0.000	[0.5756663, 1.1427850]
CC 50–85	0.303553	0.1240518	2.45	0.014	[0.0604159, 0.5466901]
AE	0.6194522	0.2046001	3.03	0.002	[0.2184433, 1.0204610]
Critical	1.086517	0.147657	7.36	0.000	[0.797115, 1.3759200]
LV function					
Moderate	0.3150652	0.1036182	3.04	0.002	[0.1119773, 0.5181530]
Poor	0.8084096	0.1498233	5.40	0.000	[0.5147614, 1.1020580]
Very poor	0.9346919	0.2917754	3.20	0.001	[0.3628227, 1.5065610]
Recent MI	0.1528943	0.136257	1.12	0.262	[−0.1141646, 0.4199531]
PA systolic pressure					
31–55 mmHg	0.1788899	0.1266713	1.41	0.158	[−0.0693812, 0.4271611]
≥55	0.3491475	0.1676641	2.08	0.037	[0.0205318, 0.6777632]
Urgency					
Urgent	0.3174673	0.1174178	2.70	0.007	[0.0873326, 0.5476020]
Emergency	0.7039121	0.1719835	4.09	0.000	[0.3668306, 1.0409940]
Salvage	1.362947	0.33706	4.04	0.000	[0.7023221, 2.0235730]
Weight of procedure					
1 non-CABG	0.0062118	0.1463574	0.04	0.966	[−0.2806434, 0.2930670]
2	0.5521478	0.1268137	4.35	0.000	[0.3035975, 0.8006980]
3+	0.9724533	0.1463969	6.64	0.000	[0.6855206, 1.2593860]
Thoracic aorta	0.6527205	0.221183	2.95	0.003	[0.2192097, 1.0862310]
Constant	−5.324537	0.1682446	−31.65	0.000	[−5.65429, −4.9947830]

NYHA: New York Heart Association; CCS: Canadian Cardiovascular Society; IDDM: insulin-dependent diabetes mellitus; ECA: extracardiac arteriopathy; CPD: chronic pulmonary dysfunction; N/M mob: neurological or musculoskeletal dysfunction severely affecting mobility; Redo: previous cardiac surgery; CC: creatinine clearance; AE: active endocarditis; Critical: critical preoperative state; LV: left ventricle; MI: myocardial infarction; PA: pulmonary artery; CABG: coronary artery bypass grafting. Weight of procedure '1 non-CABG': single major cardiac procedure which is not isolated CABG; 2: two major cardiac procedures; 3+: three or more major cardiac procedures. For age, $X_i = 1$ if patient age ≤60; X_i increases by one point per year thereafter (age 60 or less $X_i = 1$; age 61 if $X_i = 2$; age 62 if $X_i = 3$ and so on).

Source: Nashef SAM et al. *EJCTS*. 2012; 41: 1-12.

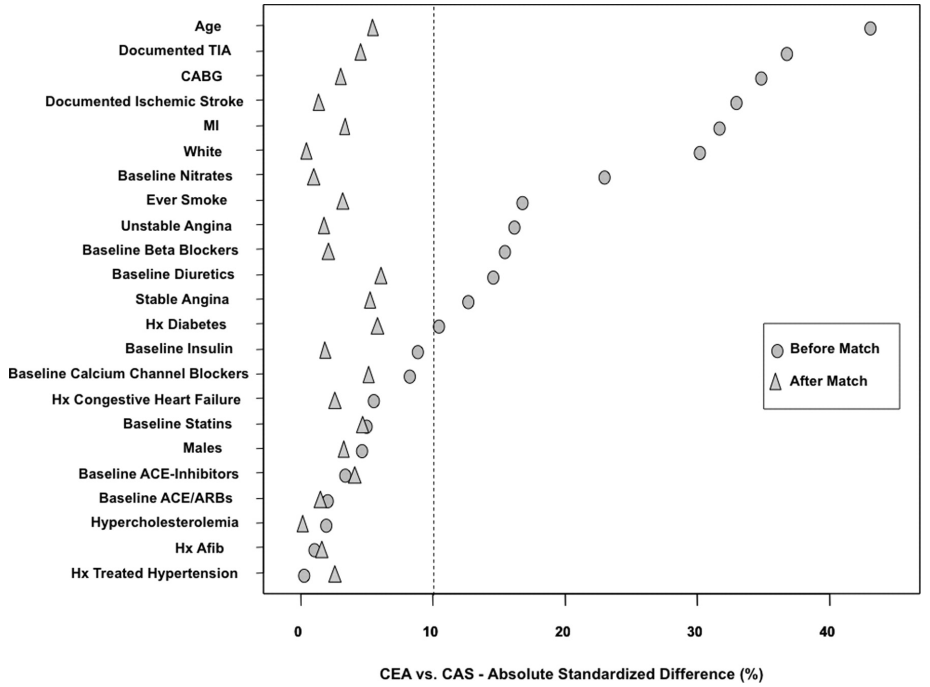
Figures *or* tables

But avoid repetition/duplication

Table 1. Baseline Characteristics of Patients With CAS Versus CEA Before and After Propensity Score Matching

	Before Propensity Score Matching			After Propensity Score Matching		
Parameter	CAS (n=1025)	CEA (n=2387)	*P*	CAS (n=836)	CEA (n=836)	*P*
Age (SD), y	68±10	72±9	<0.001	70±9	69±9	0.260
Men, n (%)	709 (69.2)	1604 (67.2)	0.242	569 (68.1)	557 (66.6)	0.531
White, n (%)	656 (64.0)	1849 (77.5)	<0.001	577 (69.0)	576 (68.9)	0.958
Hypertension, n (%)	870 (84.9)	2026 (84.9)	0.999	719 (86.0)	726 (86.8)	0.617
Diabetes mellitus, n (%)	434 (42.3)	892 (37.4)	0.006	342 (40.9)	365 (43.7)	0.255
Dyslipidemia, n (%)	820 (80.0)	1893 (79.3)	0.645	655 (78.3)	655 (78.3)	1.000

+ figure



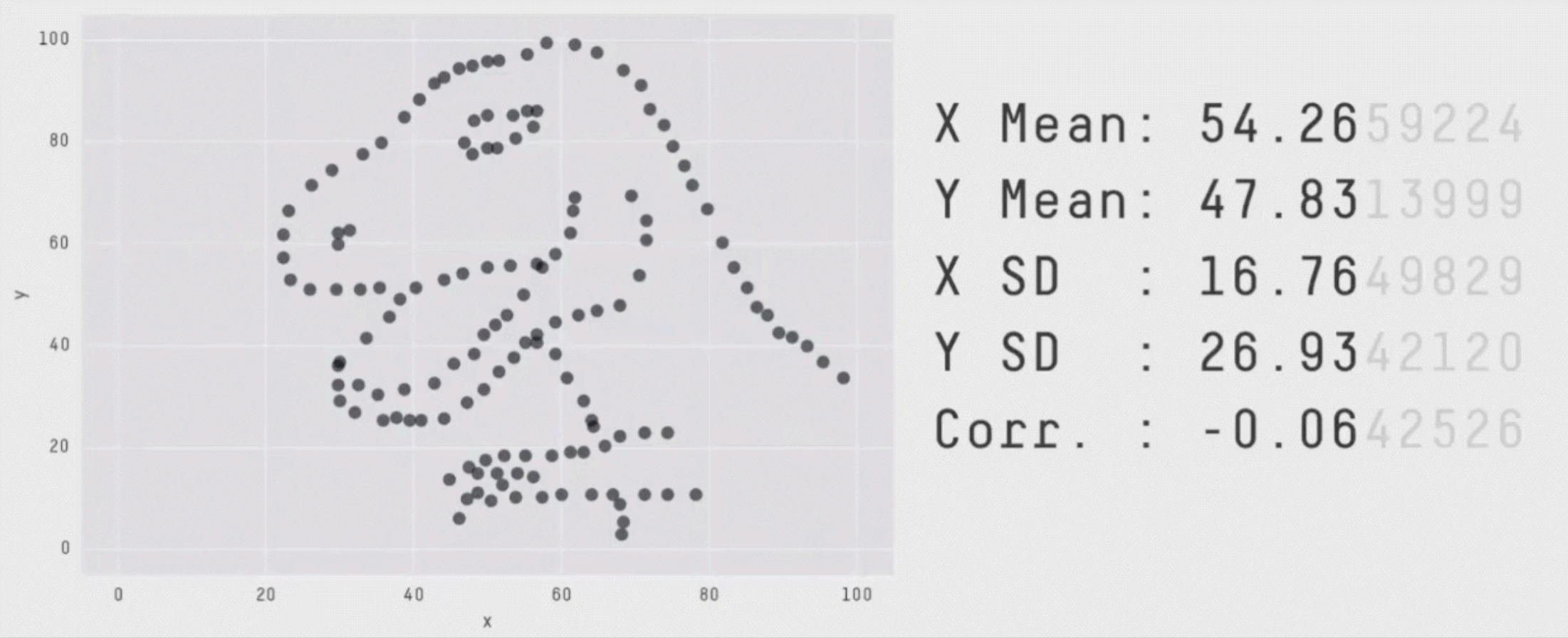
?

+ extra columns

	Δ (%): before PS matching	Δ (%): after PS matching
Age (years)	42.1	-11.0
Men	-4.3	-3.2
White	30.0	-0.2
Hypertension	0.0	2.3
Diabetes mellitus	-10.0	5.7
Dyslipidemia	1.7	0.0

Source: Bangalore et al. *Circulation*. 2010; 122: 1091-1100

Don’t trust summary statistics alone



Source: Matejka & Fitzmaurice (2017) https://www.autodeskresearch.com/publications/samestats
http://dx.doi.org/10.1145/3025453.3025912

Show all the data

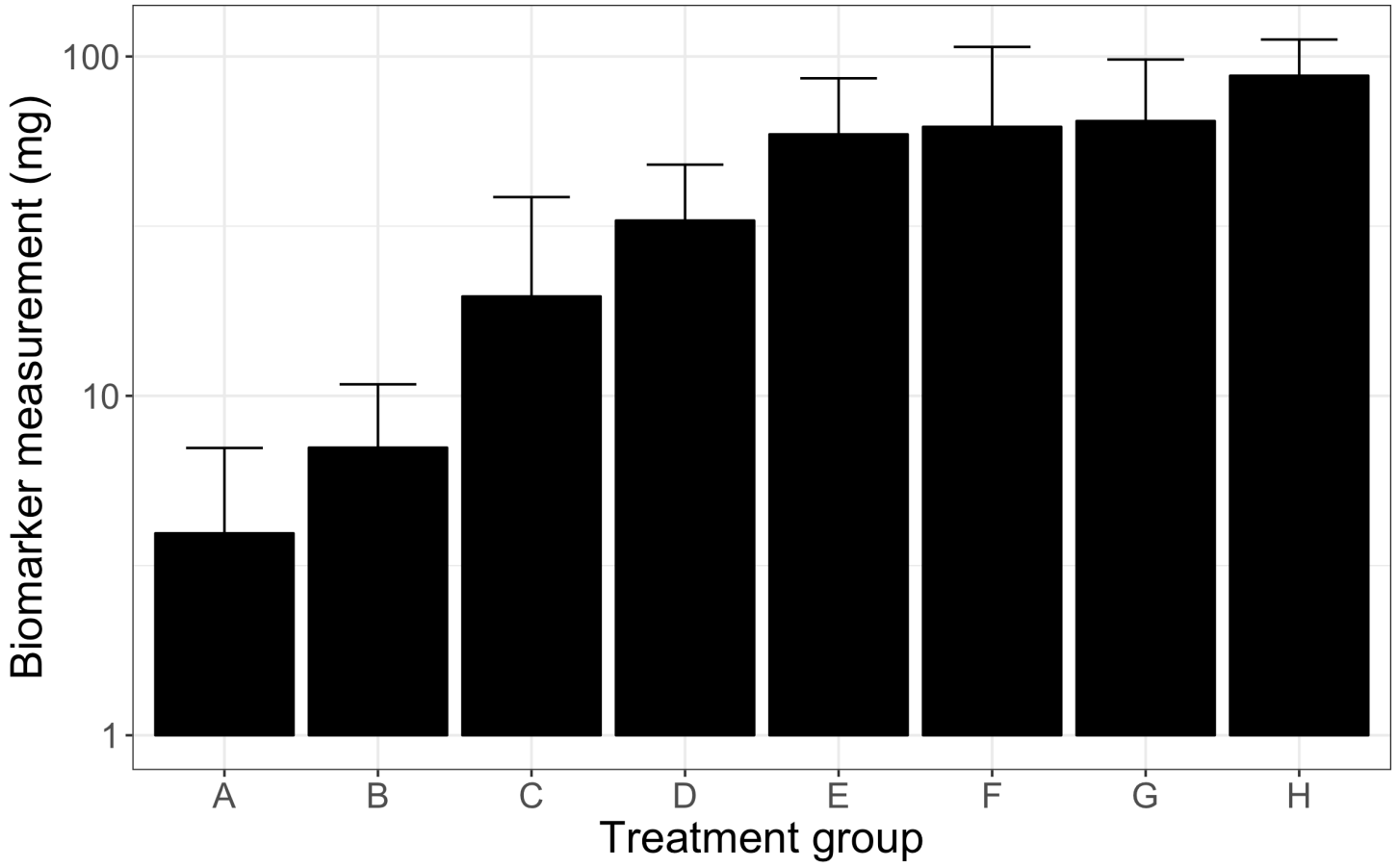
> We will ask authors, where possible, not to use bar graphs, and instead to use approaches that present full data distribution.

nature

2017

Source: http://www.nature.com/news/announcement-towards-greater-reproducibility-for-life-sciences-research-in-nature-1.22062
Nature 546, 8 (01 June 2017) doi:10.1038/546008a

Show all the data: dynamite plot



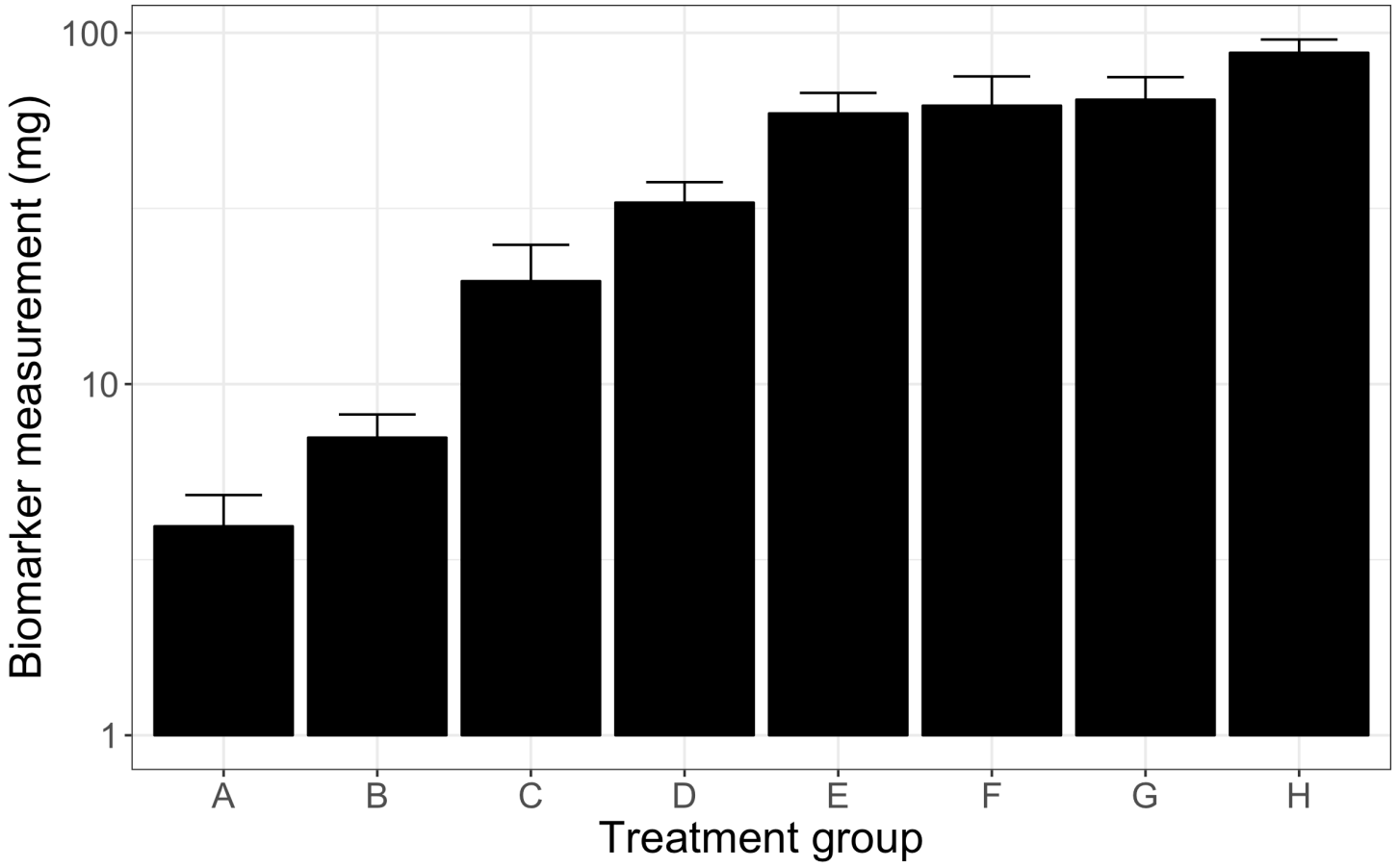
Shows:

- mean
- 1 standard deviation (SD)

Hides:

- the data
- asymmetry
- multi-modality
- lower error bar

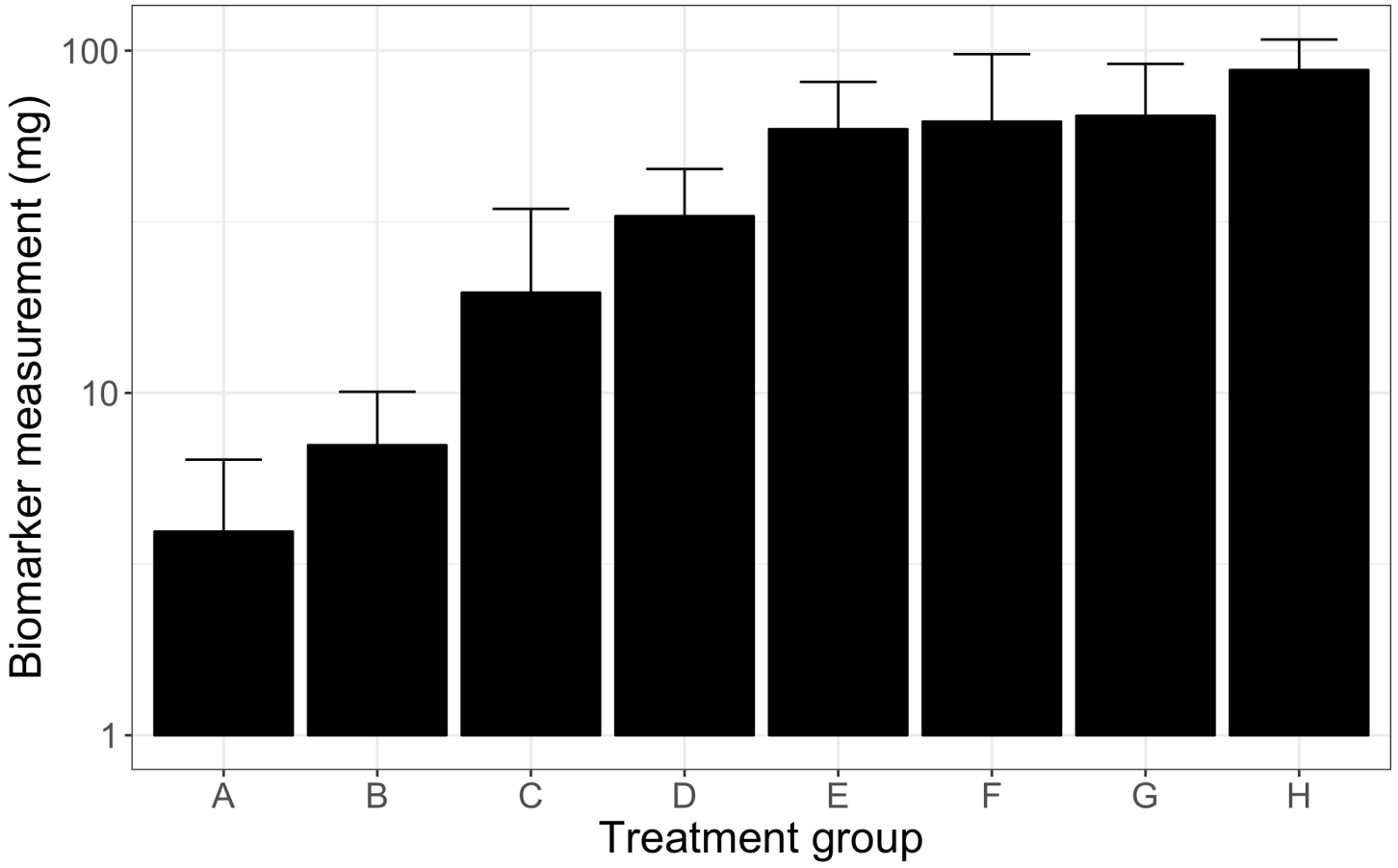
Show all the data: dynamite plot



Shows:

- mean
- 1 standard error (SEM)

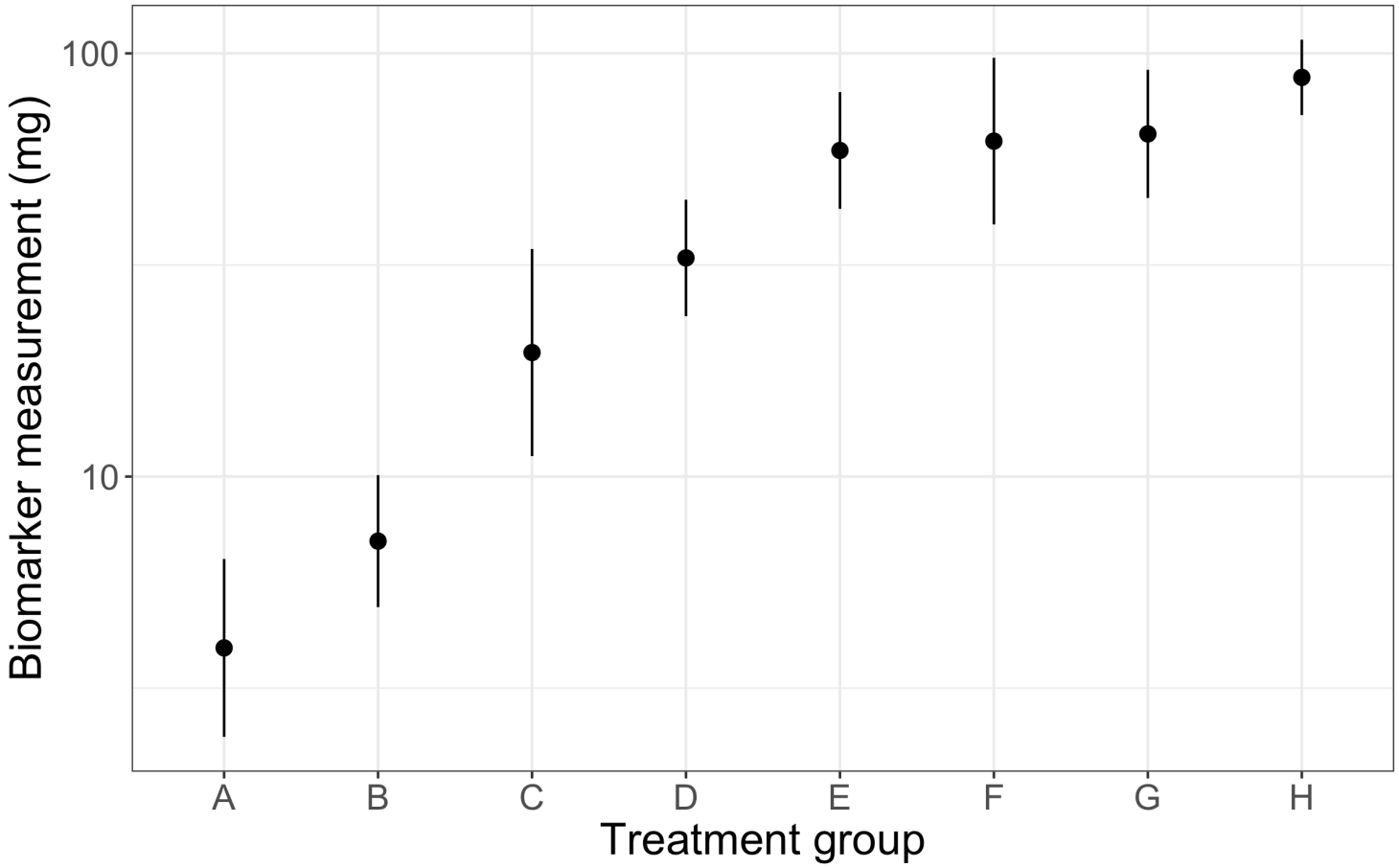
Show all the data: dynamite plot



Shows:

- mean
- 95% confidence interval (CI)

Show all the data: error bar plot

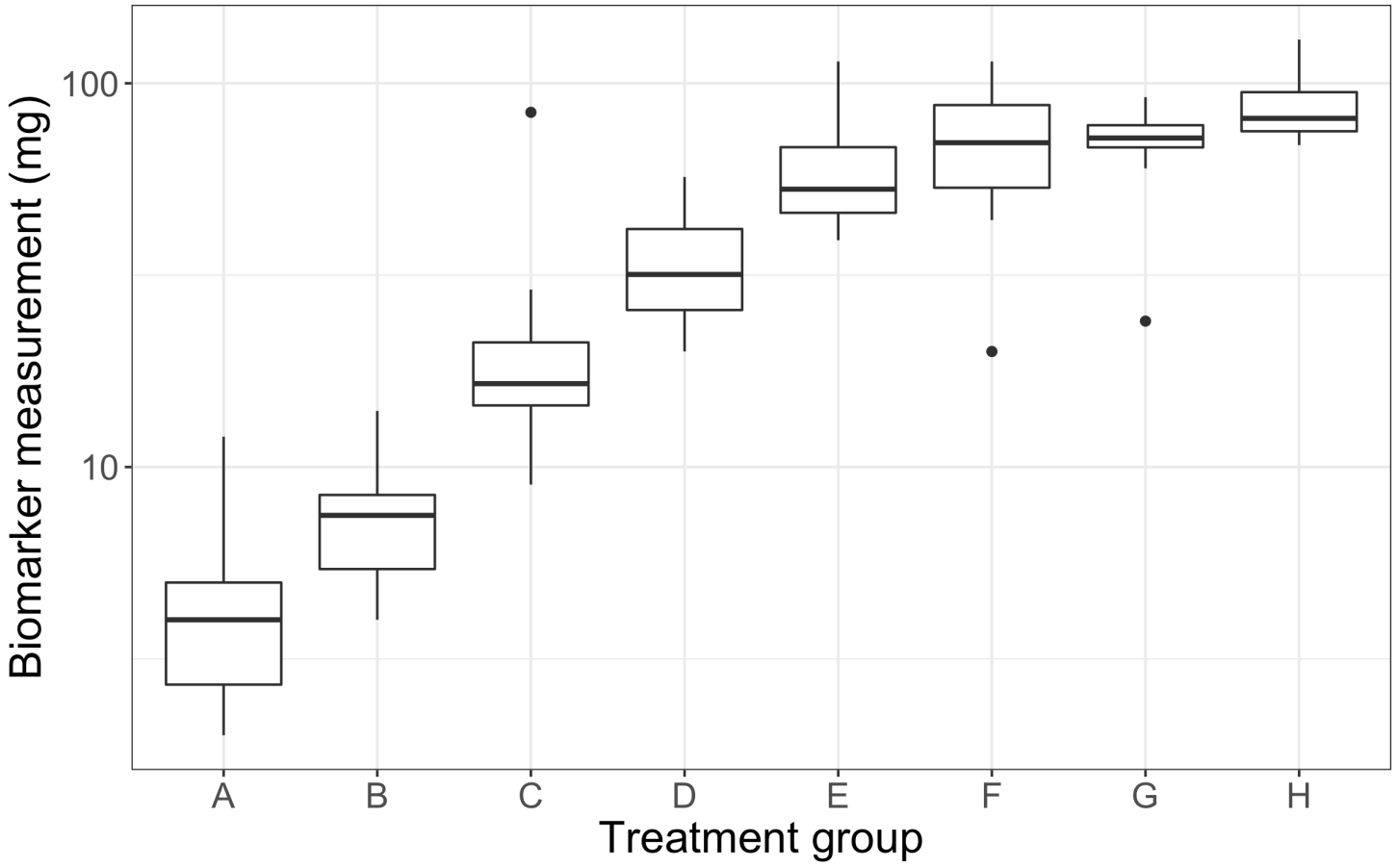


Shows:

- mean
- 95% confidence interval (CI)

A little better, but still shares a lot of limitations

Show all the data: box and whisker plot

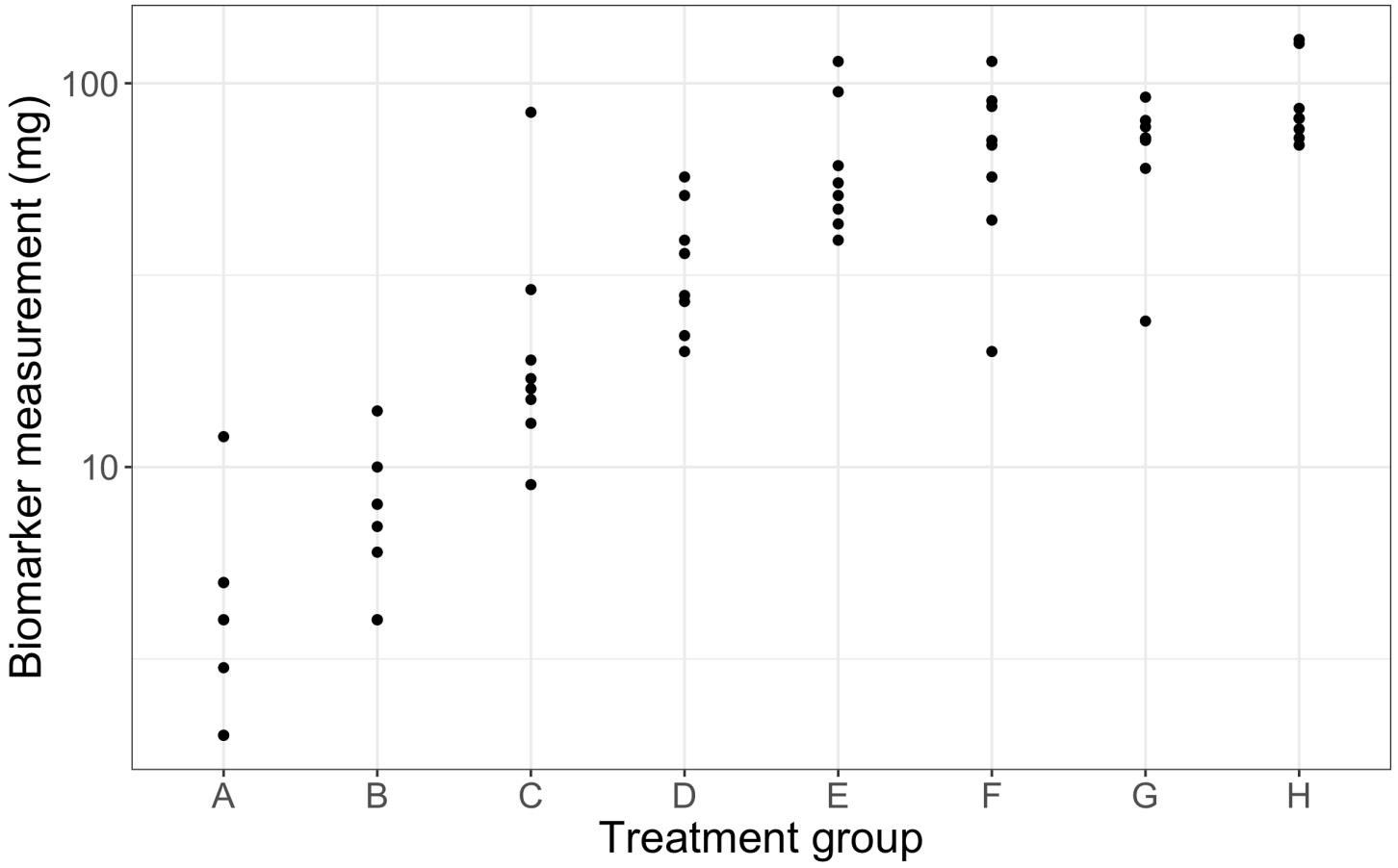


Shows:

- median
- lower & upper quartiles
- outliers
- lowest/highest values within 1.5 IQR

Up until now, my preferred choice of plot

Show all the data: dot plot



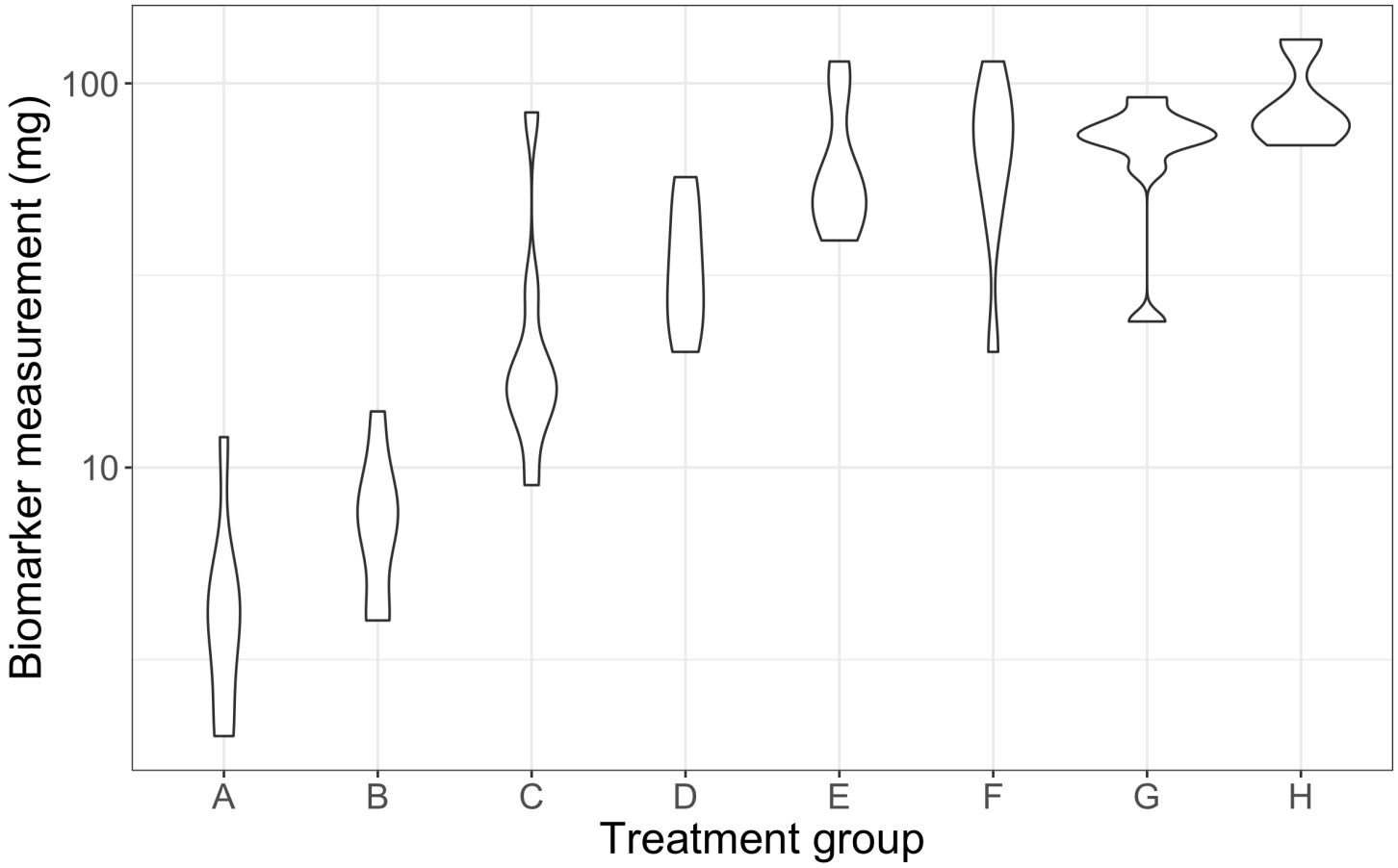
Shows:

- raw data only

Doesn't show:

- summary statistics

Show all the data: violin plot



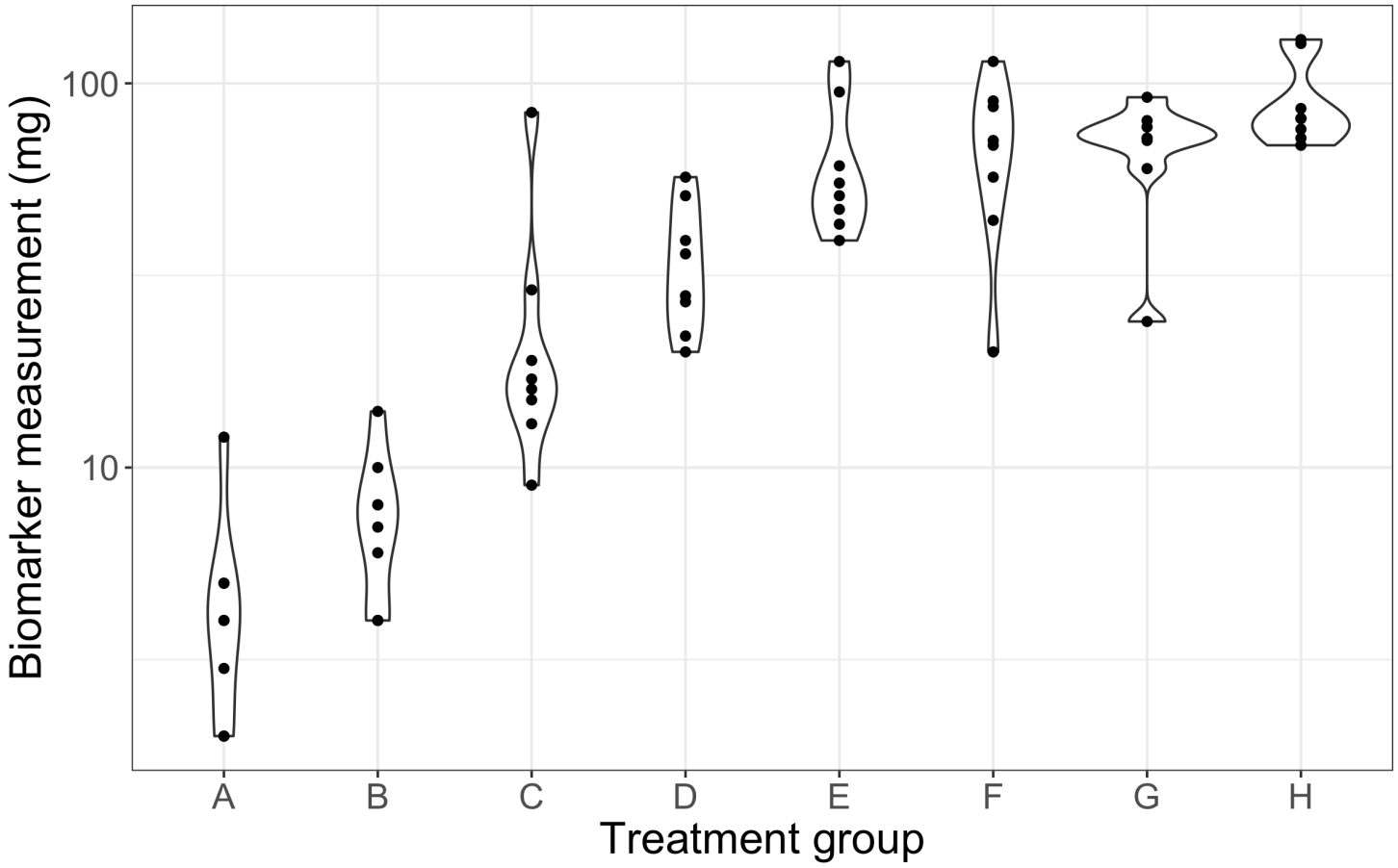
Shows:

- densities

Limitations:

- unfamiliar
- symmetry in densities arbitrary

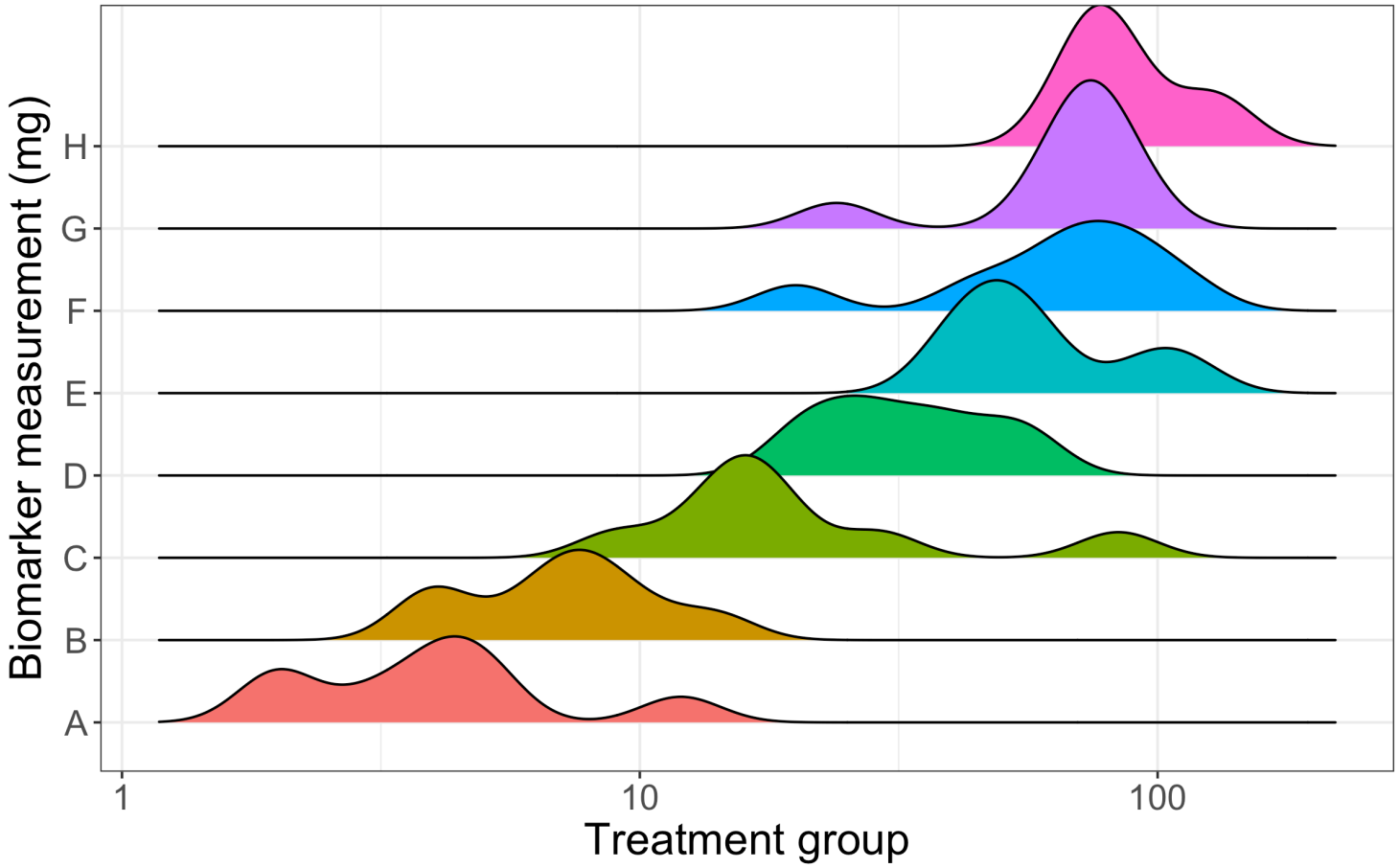
Show all the data: violin + dot plot



Shows:

- densities
- raw data

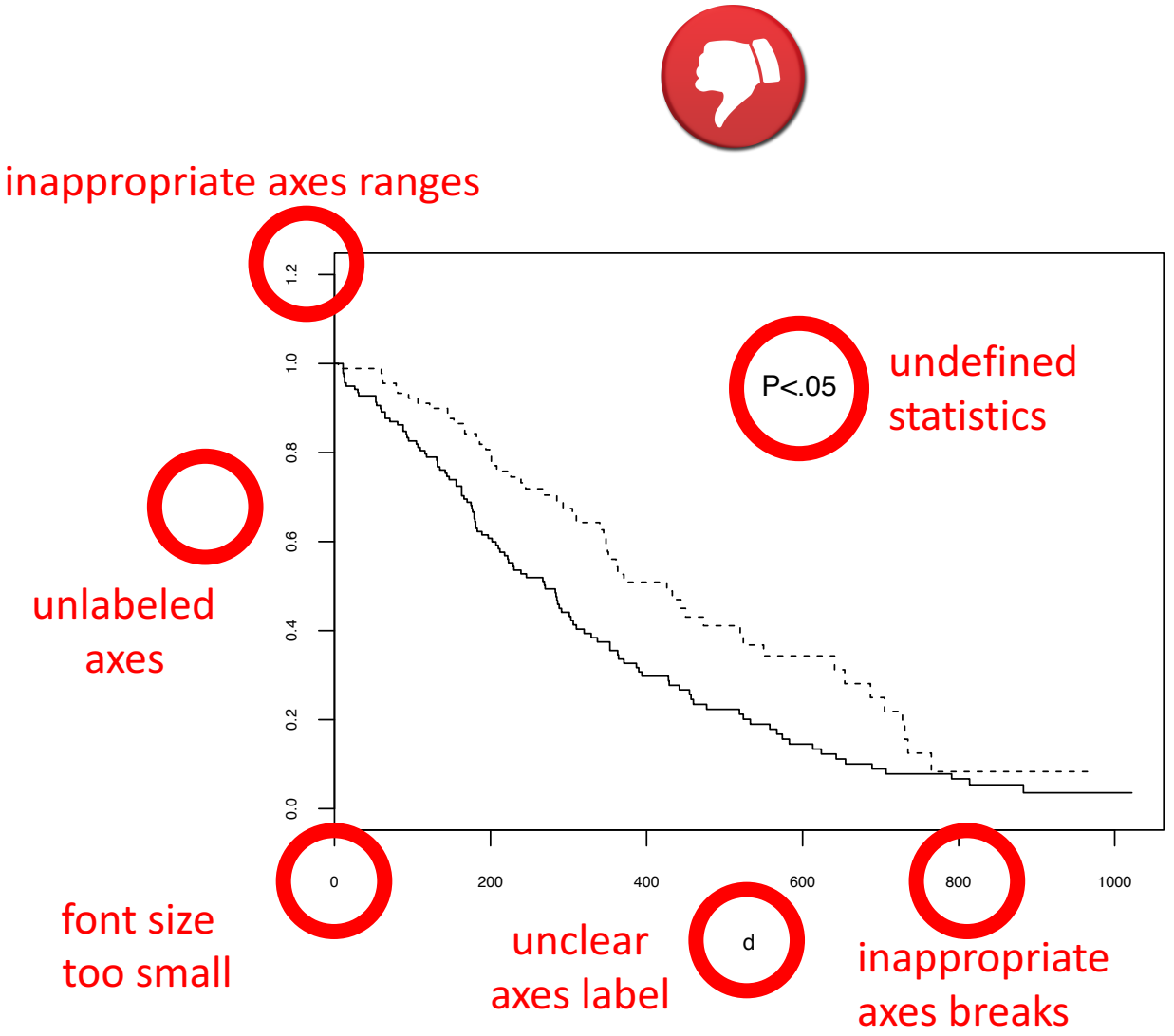
Show all the data: ridgeline plot

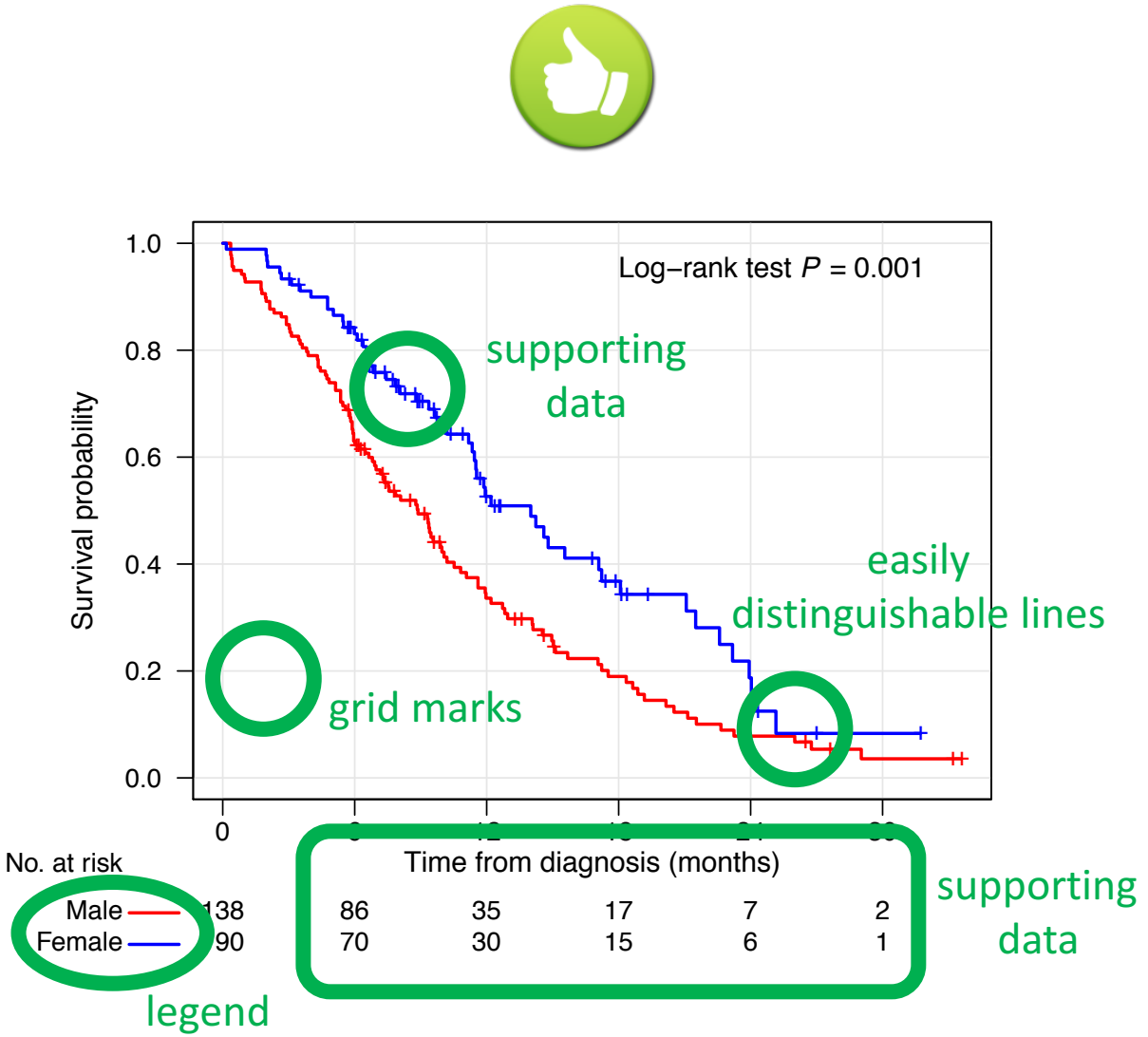


Shows:

- densities

The anatomy of a (non-)informative figure





Tables that confuse

	A (*N*=56)	B (*N*=56)
Age (years)	64.5	63.2746
Female	24 (42.8%)	32 (57.14%)
NYHA		
I	7	1
II	23	19
III	22	25
IV	3	10
Creatinine	1.2 (0.9 – 1.5)	1.6 (1.1 to 3.2)
Abnormal CRP	8 (14.3%)	28 (50.0%)

Some of the things that I comment on most frequently:

- Missing statistics (e.g. standard deviation)
- Inappropriate precisions
- Inconsistent precisions
- Percentages incorrectly calculated
- Data don’t add up
- Missing measurement units (e.g. mg/dL or µmol/L?)
- Undefined statistics
- Undefined variables
- …

Use figures to *inform*, not *confuse*

Things to (probably) avoid

3D charts

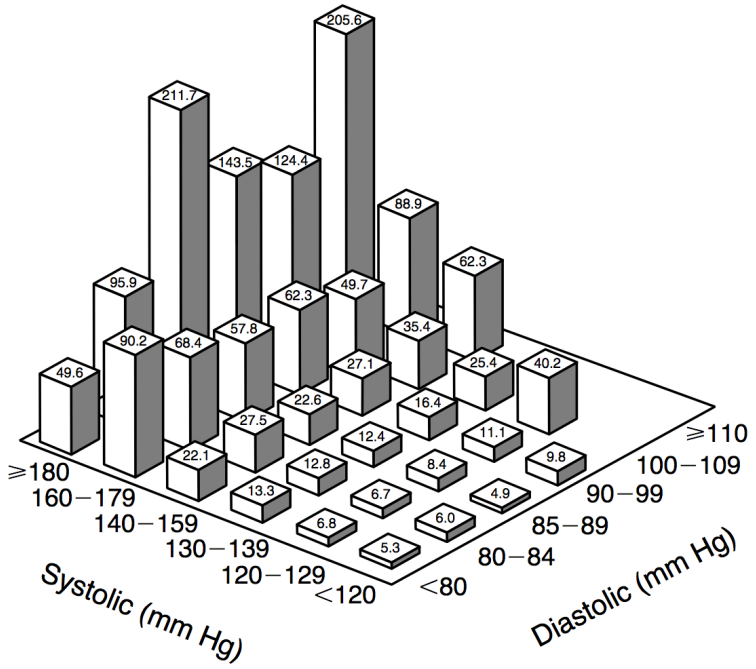
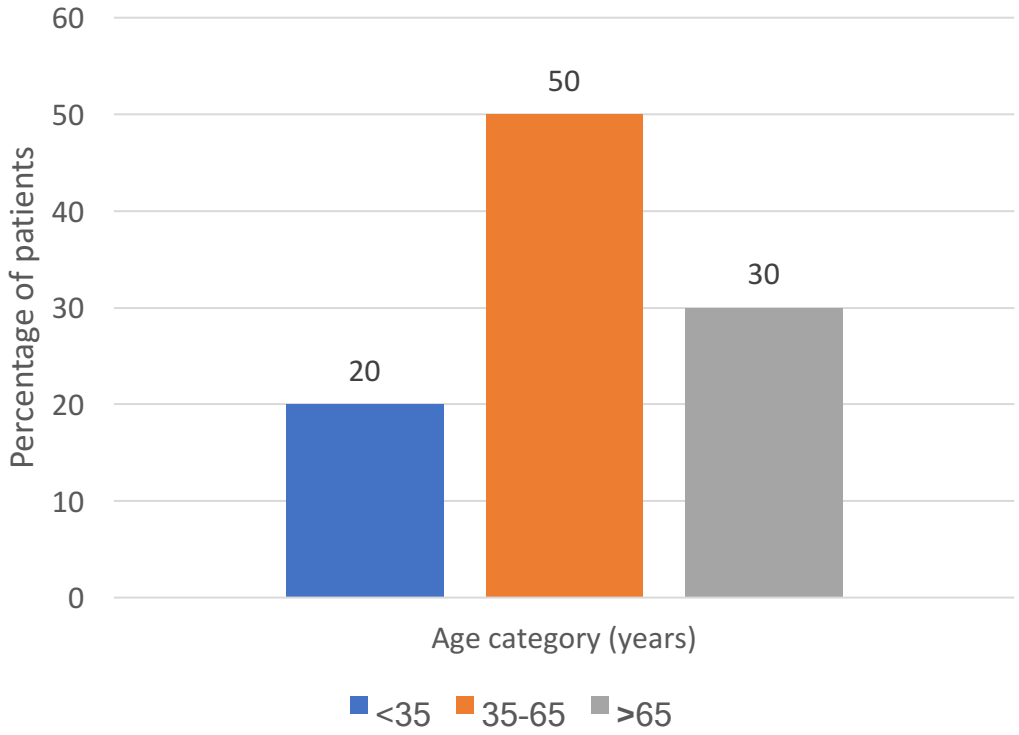


Figure 2. Age-Adjusted Rate of End-Stage Renal Disease Due to Any Cause per 100,000 Person-Years, According to Systolic and Diastolic Blood Pressure in 332,544 Men Screened for MRFIT.

- 3rd dimension adds no information
- Difficult for comparison
- Often can't read-off values

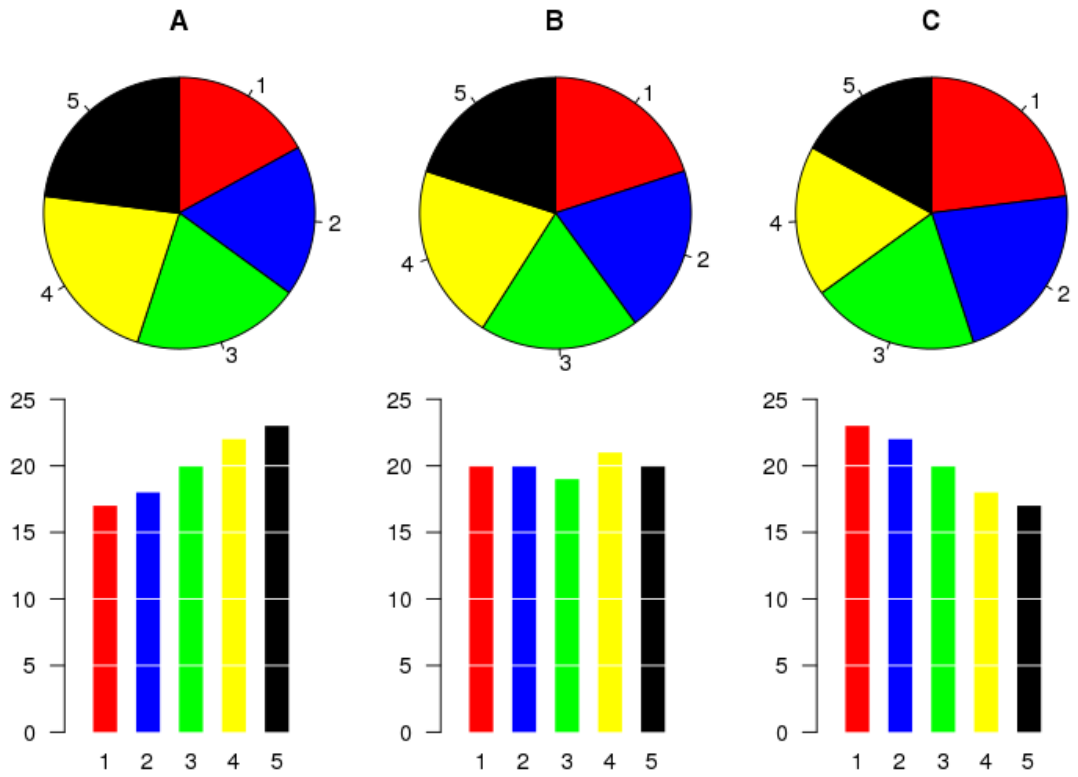
Superfluous plots



- Waste of page space
- Often repeating information in main text

Source: Klag et al. *N Engl J Med* 1996; 334:13-18

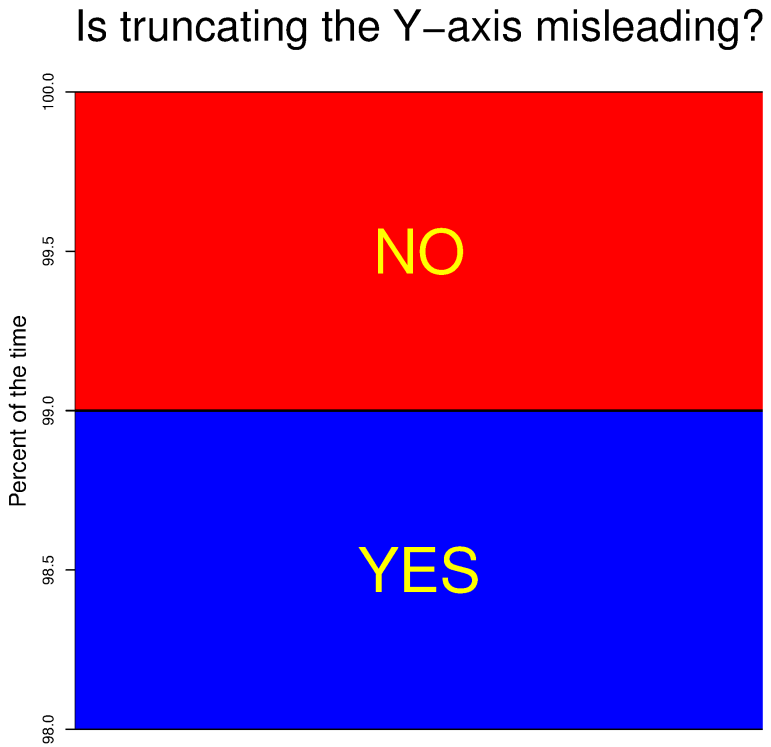
Pie charts



- Unusable for large amounts of data
- Difficult for comparison
- Can’t display trends / patterns

Source: https://en.wikipedia.org/wiki/Pie_chart

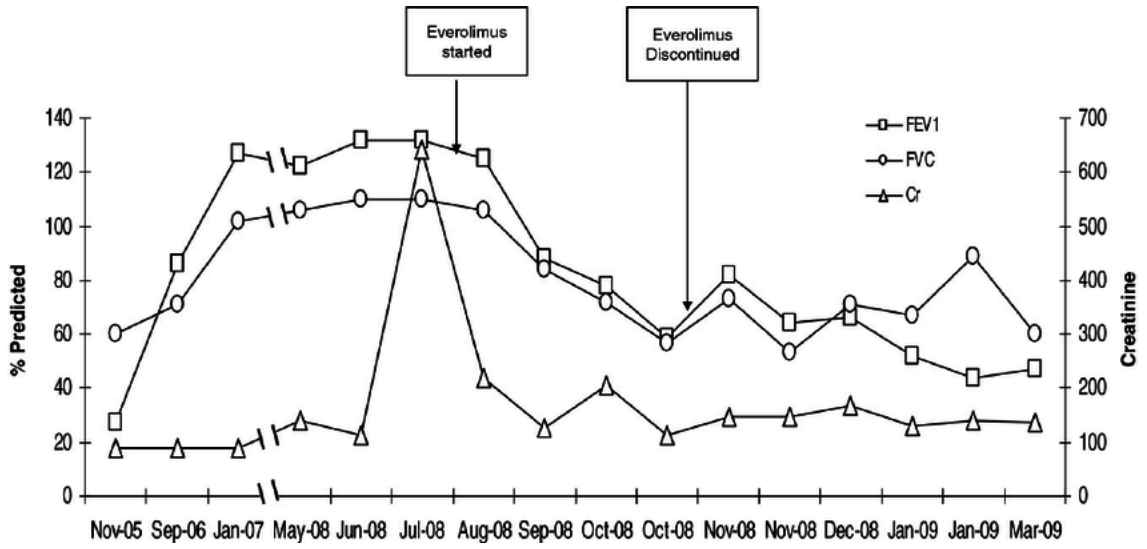
Truncated axes



- Easily misinterpreted
- Often not consistent across multiple plots

Source: http://the-geophysicist.com/lying-with-statistics

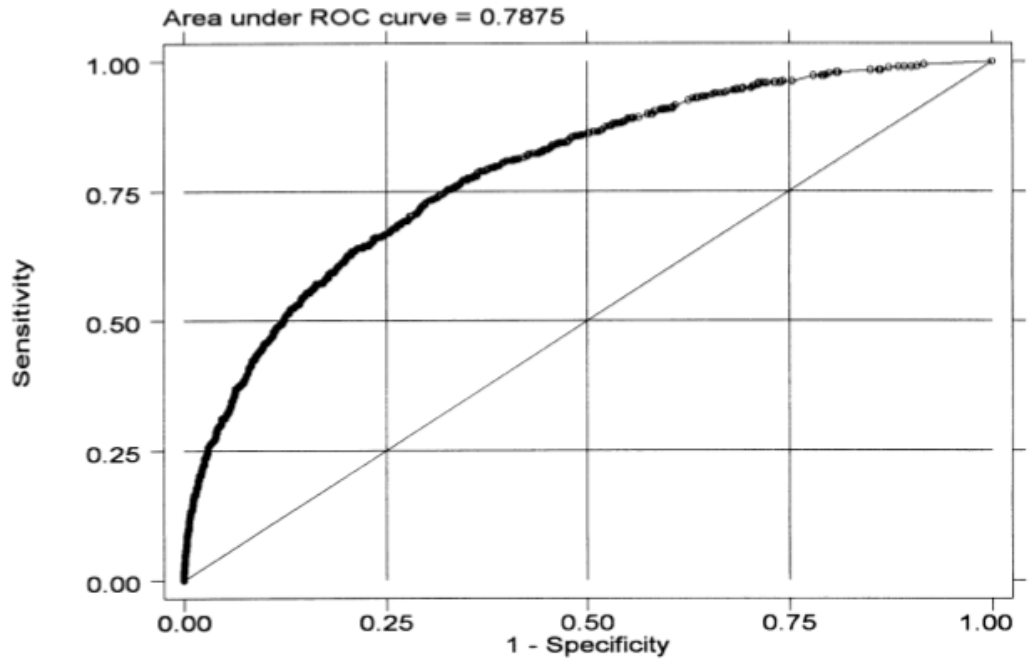
Dual y-axis graphs



- Confusing and distracting
- Often poorly labelled

Source: Keating et al. *The Annals of Thoracic Surgery*. 2011; 92: 1893-6

ROC plots



- Graphs presented often provide no extra information beyond the AUROC

Source: Nashef SAM et al. *Eur J Cardio-Thoracic Surg*. 1999;16: 9–13.

Where to get EJCTS/ICVTS specific advice

EJCTS & ICVTS Statistical and Data Reporting Guidelines

European Journal of Cardio-Thoracic Surgery 48 (2015) 180–193
doi:10.1093/ejcts/ezv168 Advance Access publication 12 May 2015

GUIDELINE

Cite this article as: Hickey GL, Dunning J, Seifert B, Sodeck G, Carr MJ, Burger HU *et al.* Statistical and data reporting guidelines for the *European Journal of Cardio-Thoracic Surgery* and the *Interactive CardioVascular and Thoracic Surgery*. Eur J Cardiothorac Surg 2015;48:180–93.

Statistical and data reporting guidelines for the *European Journal of Cardio-Thoracic Surgery* and the *Interactive CardioVascular and Thoracic Surgery*

Graeme L. Hickey[a,b,c,*], Joel Dunning[d], Burkhardt Seifert[e], Gottfried Sodeck[f], Matthew J. Carr[g], Hans Ulrich Burger[h] and Friedhelm Beyersdorf[i] on behalf of the *EJCTS* and *ICVTS* Editorial Committees

[a] Department of Epidemiology and Population Health, Institute of Infection and Global Health, University of Liverpool, The Farr Institute@HeRC, Liverpool, UK
[b] National Institute for Cardiovascular Outcomes Research (NICOR), University College London, London, UK
[c] Academic Surgery Unit, University of Manchester, Manchester Academic Health Science Centre, University Hospital of South Manchester, Manchester, UK
[d] Department of Cardiothoracic Surgery, James Cook University Hospital, Middlesbrough, UK
[e] Department of Biostatistics, Epidemiology, Biostatistics and Prevention Institute, University of Zürich, Zurich, Switzerland
[f] Department of Emergency Medicine, Medical University Vienna, Vienna, Austria
[g] University of Manchester, Institute of Brain, Behaviour and Mental Health, Manchester, UK
[h] Hoffmann-La Roche AG, Basel, Switzerland
[i] Department of Cardiovascular Surgery, Freiburg University Heart Center, Freiburg, Germany

[*] Corresponding author. Department of Epidemiology and Population Health, University of Liverpool, Institute of Infection and Global Health, The Farr Institute@HeRC, Waterhouse Building (Block F), 1-5 Brownlow Street, Liverpool L69 3GL, UK. Tel: +44-151-7958306; e-mail: graemeleehickey@gmail.com (G. Hickey).

Received 27 March 2015; accepted 2 April 2015

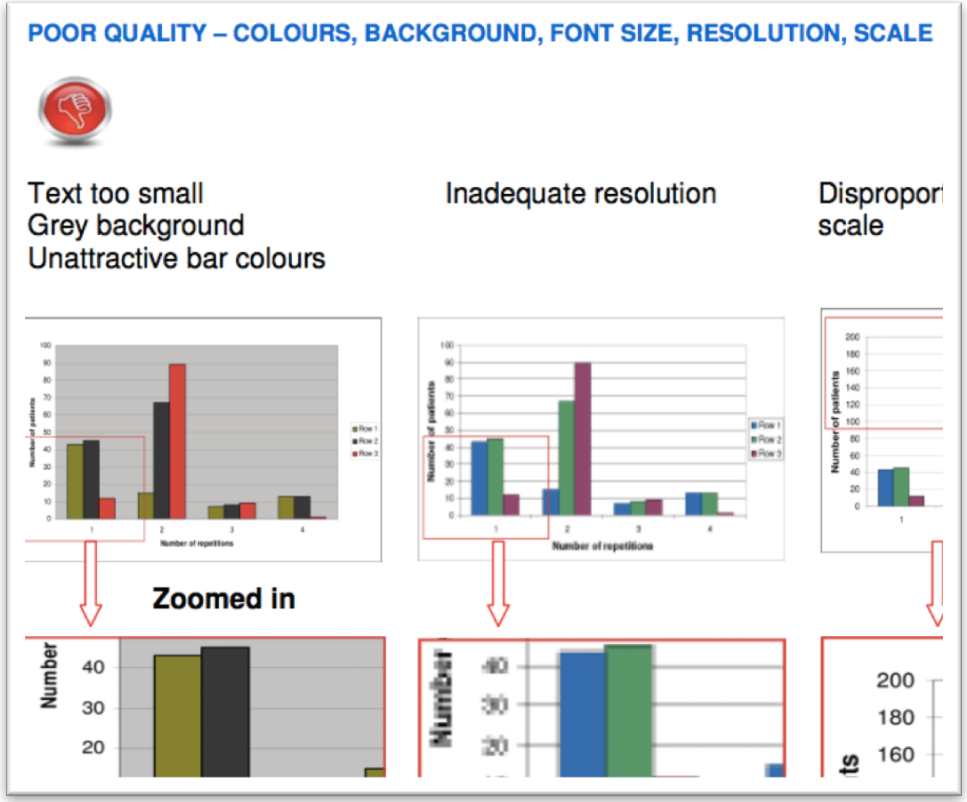
Abstract

As part of the peer review process for the *European Journal of Cardio-Thoracic Surgery* (EJCTS) and the *Interactive CardioVascular and Thoracic Surgery* (ICVTS), a statistician reviews any manuscript that includes a statistical analysis. To facilitate authors considering submitting a manuscript and to make it clearer about the expectations of the statistical reviewers, we present up-to-date guidelines for authors on statistical and data reporting specifically in these journals. The number of statistical methods used in the cardiothoracic literature is vast, as are the ways in which data are presented. Therefore, we narrow the scope of these guidelines to cover the most common applications submitted to the *EJCTS* and *ICVTS*, focusing in particular on those that the statistical reviewers most frequently comment on.

Keywords: Guidelines • Statistics • Data • Reporting • Peer-review

Source: Hickey et al. *Eur J Cardiothorac Surg* 2015;48:180–93.

EJCTS/ICVTS Instructions for Authors webpage

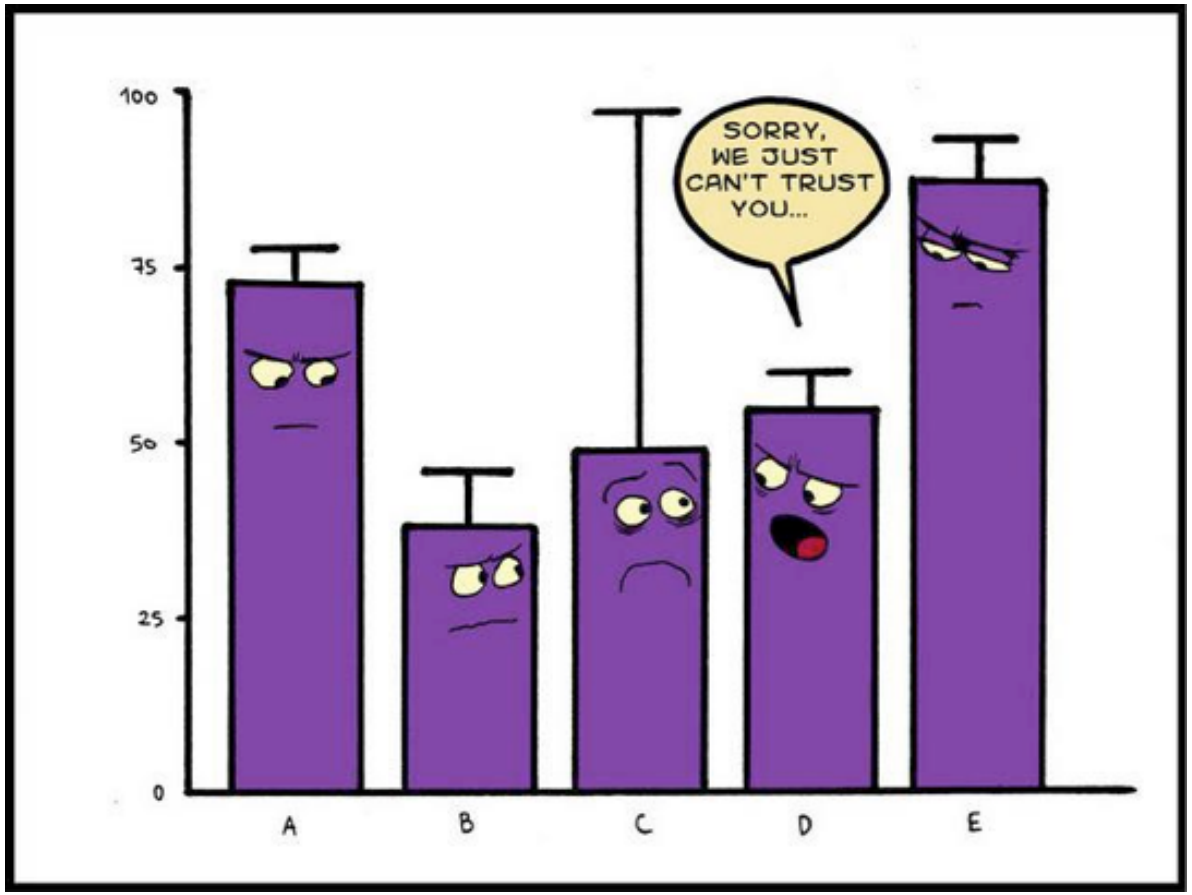


Source: https://academic.oup.com/ejcts/pages/Manuscript_Instructions

Conclusions

- Tables and figures should (ideally) be:
  - Used only if required
  - Self-contained (i.e. can be read standalone)
  - Easy to interpret
  - Clearly labelled (legends, column titles, etc.)
  - Neatly presented (high quality figures, legible font sizes, etc.)
- Figure + Table legends are effective constructs for conveying extra information that facilitates interpretation

- I always look at the figures and tables first when reviewing a paper



Thank you for listening…

any questions?

Slides available (shortly) from: www.glhickey.com