

Measurement

Doug Altman

**EQUATOR Network, Centre for Statistics in Medicine,
NDORMS, University of Oxford**

**EQUATOR – OUCAGS training course
4 October 2014**

Measurements

- **Measurements are critical in clinical practice and in research**
- **Eligibility criteria**
- **Exposures**
- **Outcomes**
- **Choice of outcome affects whole study**



Attributes of a measurement

- **How easy are measurements to make?**
- **How important is it to have measurements?**
- **How important is it to have good measurements?**
- **What is a “good” measurement?**
 - How do we know when we have one?
- **What do we do with the measurements?**



What do we measure?

... and how easy is it to measure?

- **Weight**
- **Height**
- **Length of newborn baby**
- **Blood pressure**
- **Serum cholesterol**
- **Fetal abdominal circumference**
- **Brain volume**
- **Stenosis of an artery**
- **Time to walk 100 metres**
- **Depression**
- **Fatigue**
- **Pain**
- **Diet**
- **Quality of life**

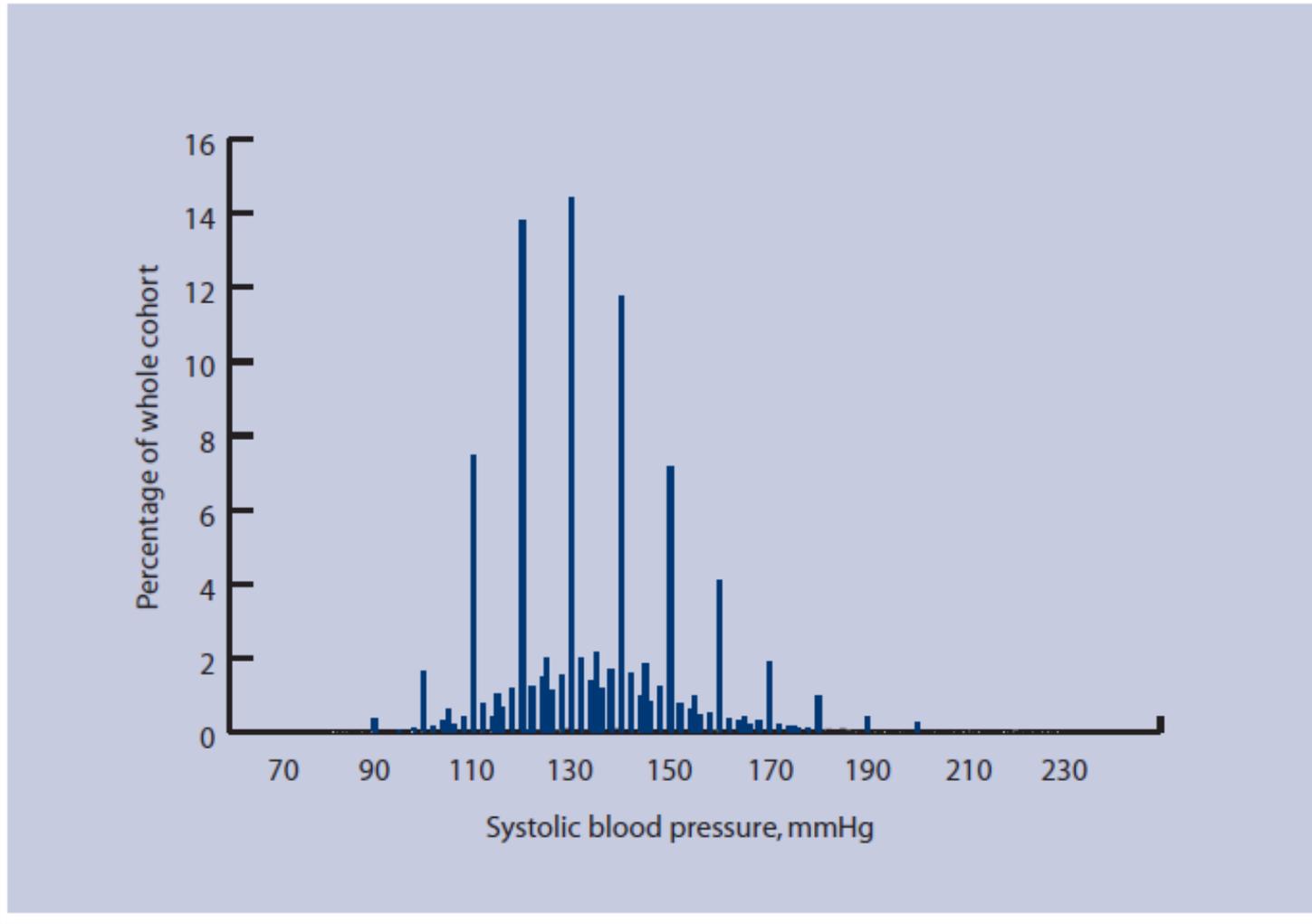


Factors that can affect measurements

- **Quantity being measured varies**
 - e.g. blood pressure, ejection fraction
- **The operator**
 - Skill/experience
 - Technique
- **Equipment**
 - Within-machine error
 - Between-machine variation
- **Ambient factors, e.g. temperature**
- **Patient's behaviour** (e.g. recent diet, exercise)
- **Digit preference**
- **Knowledge** (e.g. clinical history, previous measurement)



Figure 1. Distribution of (a) systolic and (b) diastolic blood pressure in 23 676 patients.



[Broad et al, *Br J Gen Pract* 2007]



Zero end-digit preference in recorded blood pressure and its impact on classification of patients for pharmacologic management in primary care – PREDICT-CVD-6

Joanna Broad, Sue Wells, Roger Marshall and Rod Jackson

Results

Zero end-digits were recorded in 64% of systolic and 62% of diastolic blood pressures. When eligibility for drug treatment was based only on a Framingham 10-year CVD risk threshold of 20% or more, rounding misclassified one in 41 of all those patients subject to this error. Under the two guidelines which use different combinations of CVD risk and blood pressure thresholds, one in 19 would be misclassified under JBS2 and one in 12 under the BHS-IV guidelines mostly towards increased treatment.

Conclusion

Zero end-digit preference significantly increases a patient's likelihood of being classified as eligible for drug treatment. Guidelines that base treatment decisions primarily on absolute CVD risk are less susceptible to these errors.

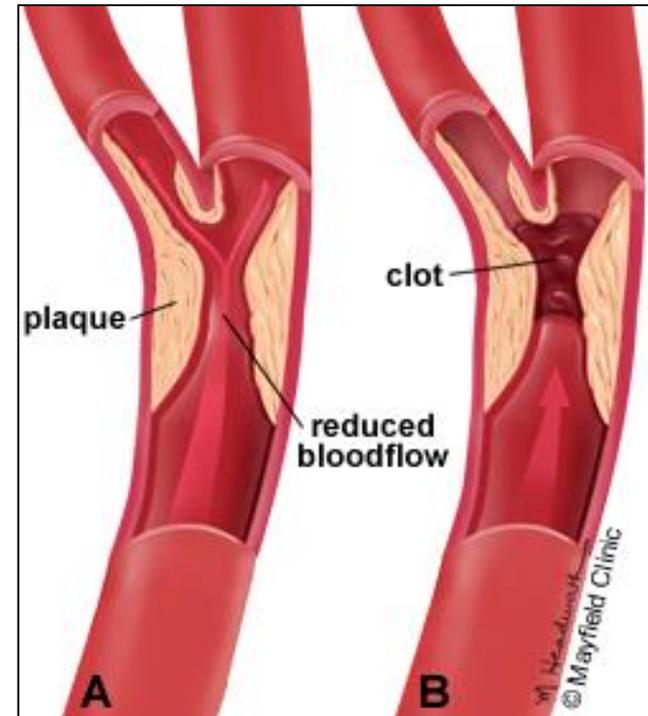
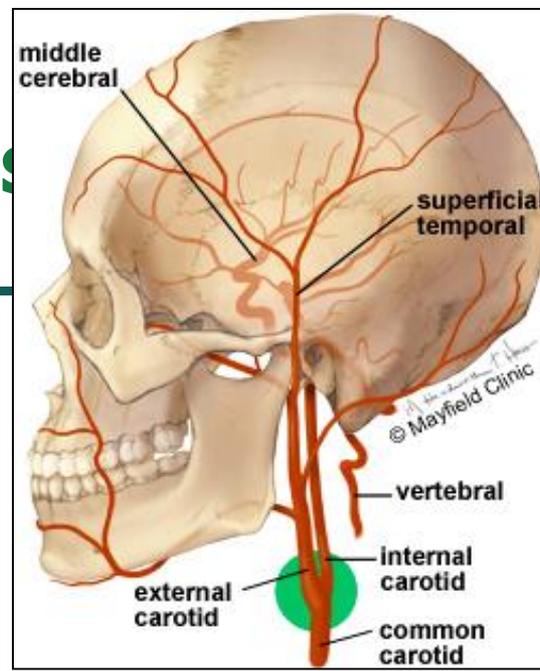


-
- **“A visually estimated stenosis of 70% or greater is considered hemodynamically significant, and is a generally accepted indication for percutaneous intervention. Unfortunately, visual estimation of the severity of stenosis has a high degree of interobserver variability and compares poorly with quantitative methods of lesion assessment.”**

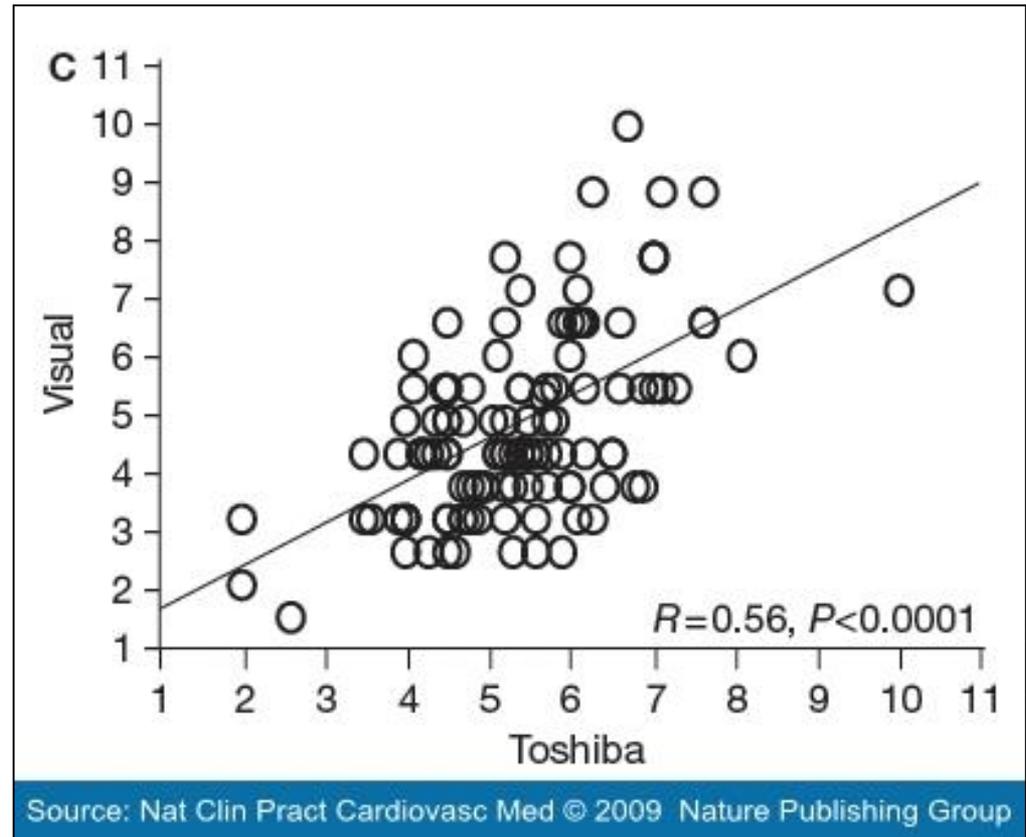
[http://www.medscape.org/viewarticle/588519_4]



Artery s



Visual estimation of peripheral arterial angiographic stenosis compared with quantitative angiographic stenosis (Toshiba)



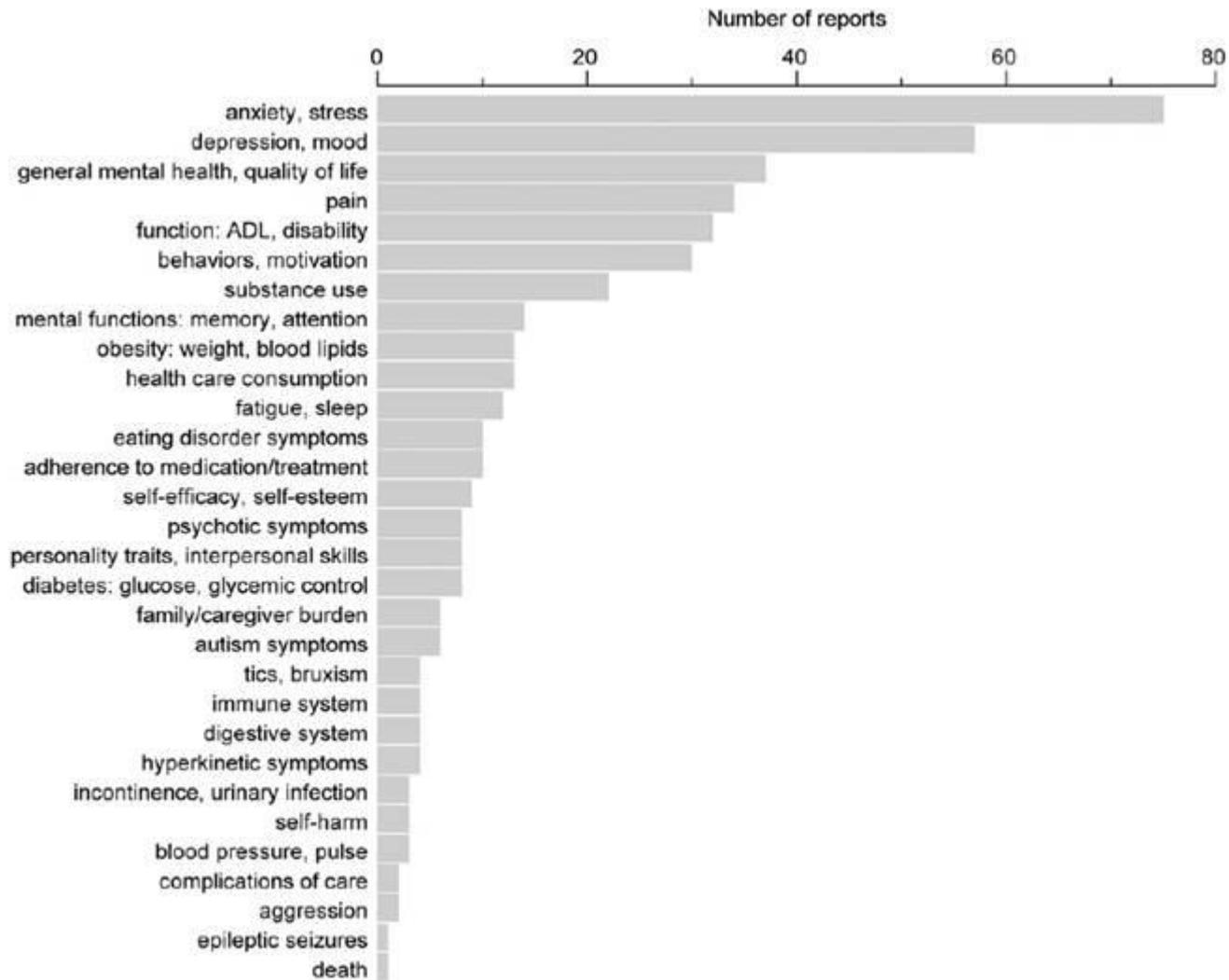


Fig. 3. Categories of outcomes in randomized controlled trials of psychological interventions published in 2010. Note: The labels serve as representative categories and include other similar outcomes.

Arnberg et al, *J Psychosom Res* 2013



What does it all mean?

- **Measurements are not necessarily true**
- **They depend not only on the thing being measured but also on:**
 - the observer
 - The equipment
 - chance, measurement error
 - mistakes
- **Observers must be trained**
- **“Data” = “those which are given”, not “that which is true”**



Validity of measurement scales

- **The ideal measurement instrument:**
 - measures what we want it to measure
 - measures with low bias
 - measures with low random error
- **How do we know it does this?**
- **Repeat measurements will vary**
- **Measurements by different instruments will vary**
- **Measurements by different observers will vary**



% body area covered by eczema

Six observers, six participants:

Observer	Participant					
	1	2	3	4	5	6
Nurse 1	5	63	4.5	18.5	3.5	2
Nurse 2	65	40	10	15	16	6
Registrar 1	30	60	5	20	10	2
Registrar 2	45.9	100	4.2	62.4	28.5	2.8
Consultant 1	7	13	0.3	5	6.5	0.4
Consultant 2	9.5	75	4.4	3	12.8	13.7

What are the problems here?

What could we do?



Types of measurement

- **Real measurements (by humans or machines)**
 - Blood pressure
 - Biochemistry, e.g. serum cholesterol
 - Fetal head circumference (ultrasound)
- **Subjective measurements**
 - Pain (visual analogue scale)
 - Scales for quality of life, depression, etc (questionnaires)
- **Surrogate measurements**
 - Easier and cheaper to measure than the real thing
 - blood pressure and lipids rather than stroke, MI, death
 - dietary recall vs diary vs weighed food intake



Low Back Disability Questionnaire (Oswestry)

Please answer every section and select in each section only the ONE CHOICE which applies to you.

We realize that you may consider that two of the statements in any one section relate to you, but please just choose the one which most closely describes your condition.

Section 1 - Pain Intensity

The pain comes and goes and is very mild.

The pain is mild and does not vary much.

The pain comes and goes and is moderate.

The pain is moderate and does not vary much.

The pain comes and goes and is very severe.

The pain is severe and does not vary much.

Section 2 - Personal Care

I would not have to change my way of washing or dressing in order to avoid pain.

I do not normally change my way of washing or dressing even though it causes some pain.

Washing and dressing increases the pain, but I manage not to change my way of doing it.

Washing and dressing increases the pain and I find it necessary to change my way of doing it.

Because of the pain, I am unable to do some washing and dressing without help.

Because of the pain, I am unable to do any washing and dressing without help.

Next



Strategies for reducing error

- 1. Standardising the measurement methods**
- 2. Training and certifying the observer**
- 3. Refining the instrument**
- 4. Repeating the measurement**
- 5. Blinding to previous measurements or other variables**
- 6. Blinding to study hypothesis**



Measuring exposure

- **Smoking / Diet / Radiation exposure / ...**
- **Current or lifetime?**
- **Use objective records rather than subjective measures if possible**
- **Standardise questionnaire**
- **Use measures of exposure taken before the outcome occurred**
 - Or blind the investigator to the outcome status of the patient
- **Don't categorise continuous measures**



Measuring outcome

- **Choice of outcome variable**
 - Affects sample size
- **Keep continuous variables as continuous**
 - BMI rather than obesity (BMI>30)
 - Change in pain score rather than whether 50% improvement
- **Trade-off between hard measurements (e.g. biochemical measurements) and softer patient reported outcomes affecting quality of life**



Too much choice of outcomes

**Review of 10000 RCTs in schizophrenia (≤ 2008)
[Miyar and Adams, *Schizophrenia Bull* 2012]**

- **1940 different interventions**
- **2194 different rating scales!**



The logo for the COMET Initiative features the word "COMET" in large, light blue, sans-serif capital letters. The letter "O" is replaced by a realistic image of the Earth. Below the "O" is a small, stylized satellite or probe with a white body and a blue circular component. Below "COMET" is the word "INITIATIVE" in a smaller, teal, sans-serif font.

COMET INITIATIVE

Core Outcome Measures in Effectiveness Trials

www.comet-initiative.org

