

Contents

- Preliminaries
- 0: Introduction
- 1: Background & Basic Concepts
- 2: Basic Trial analysis
- 3: Randomization
- 4: Protocol Deviations
- 5: Size of the Trial
- 6: Multiplicity & Interim Analysis
- 7: Crossover Trials
- 8: Combining Trials
- 9: Binary Response Data
- 10: Comparing Methods of Measurement

319

- ### Comparing Methods of Measurement
- Key reference is Martin Bland's page <http://www-users.york.ac.uk/~mb55/>
 - 2 methods used to measure something
 - ♦ Continuous
 - e.g. two instruments,
 - two technique
 - ♦ Categorical
 - two observers rating some feature on a scale
 - **How comparable are the instruments?**
 - (or observers or techniques or....)
- 320

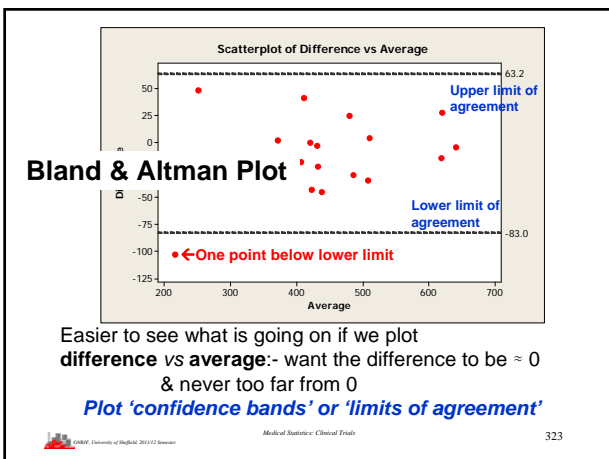
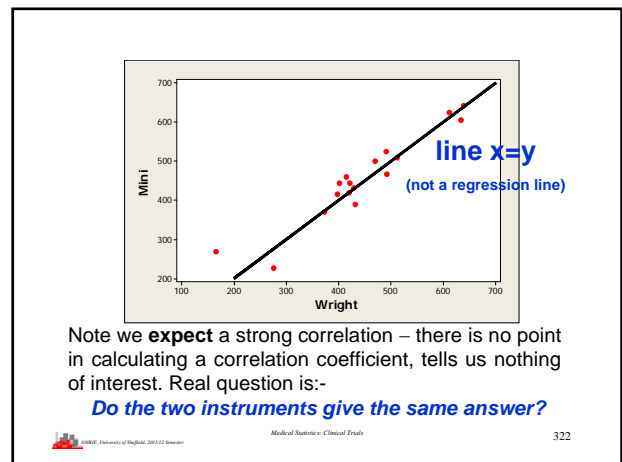
Continuous:-

How comparable are these two instruments?

Subject	Wright	Mini	Mean	Difference
1	490	525	507.5	-35
2	397	415	406.0	-18
3	512	508	510.0	4
4	401	444	422.5	-43
5	470	500	485.0	-30
6	611	625	618.0	-14
7	415	480	437.5	-45
8	431	390	410.5	41
9	638	642	640.0	-4
10	429	432	430.5	-3
11	420	420	420.0	0
12	633	605	619.0	28
13	275	227	251.0	48
14	492	467	479.5	25
15	165	268	216.5	-103
16	372	370	371.0	2
17	421	443	432.0	-22

Comparison of two methods of measuring PEFR (from Bland, 2000)

321



Discrete:- The Kappa Statistic

- Example:- **Observed categories**

First classification	Second classification		Total
	Normalizer	Non-normalizer	
Normalizer	76	17	93
Non-normalizer	39	47	86
Total	115	64	179

Expected agreements

First classification	Second classification		Total
	Normalizer	Non-normalizer	
Normalizer	$59.7 = 93 \times 115 / 179$		93
Non-normalizer		$30.7 = 86 \times 64 / 179$	86
Total	115	64	179

324



- Kappa measure of agreement defined:-

$$\kappa = \frac{A_{\text{obs}} - A_{\text{exp}}}{1 - A_{\text{exp}}}$$

- ♦ where A_{obs} and A_{exp} are observed & expected proportions of agreements
- ♦ $A_{\text{obs}} = (76+47)/179 = 0.687$
- ♦ $A_{\text{exp}} = (59.7+30.7)/179 = 0.505$
- ♦ $\kappa = (0.687 - 0.505)/(1 - 0.505) = 0.37$
- ♦ Interpretation?
 - (only 'informal' not *statistical*)

- ♦ $\kappa > 0.75$:— excellent agreement
- ♦ $0.4 < \kappa < 0.75$:— fair to good agreement
- ♦ $\kappa < 0.4$:— moderate or poor agreement

- Note: not of any interest to calculate a chi-squared value for tables of agreements since we **expect** categories to be closely related
- Generalizations to several categories & several observers possible

