Method Comparison Studies with R

Kevin O'Brien

Method Comparison Studies

- The problem of assessing the **agreement** between two or more methods of measurement is ubiquitous in scientific research, particularly with clinical sciences, and is commonly referred to as a ‘method comparison study’.
  - “Do two methods of measurement agree statistically?".
  - “Can the two methods be used interchangeably?"
- Published examples of method comparison studies can be found in disciplines as diverse as Pharmacology, Anaesthesia, and cardiac imaging methods.

Accuracy and Precision

- A method of measurement should ideally be both accurate and precise.
- Barnhart et al [7] describes agreement as being a broader term that contains both of those qualities. An accurate measurement method will give results close to the unknown 'true value'.
- The precision of a method is indicated by how tightly measurements obtained under identical conditions are distributed around their mean measurement value.
- A precise and accurate method will yield results consistently close to the true value.

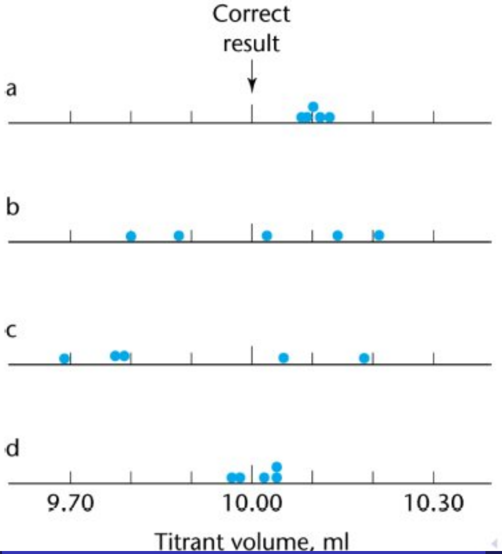
Accuracy and Precision

- Of course a method may be accurate, but not precise, if the average of its measurements is close to the true value, but those measurements are highly dispersed.
- Conversely a method that is not accurate may be quite precise, as it consistently indicates the same level of inaccuracy.
- The tendency of a method of measurement to consistently give results above or below the true value is a source of systematic bias.
- The smaller the systematic bias, the greater the accuracy of the method.

- To illustrate the point, let us consider a real experiment.
- Four students (A-D) each perform an analysis in which exactly 10.00 *ml* of exactly 0.1 M sodium hydroxide is titrated with exactly 0.1 NI hydrochloric acid.
- Each student performs five replicate titrations, with the results shown in the next slide.

Student	Results (ml)					Comment
A	10.08	10.11	10.09	10.10	10.12	Precise, biased
B	9.88	10.14	10.02	9.80	10.21	Imprecise unbiased
C	10.19	9.79	9.69	10.05	9.78	Imprecise, biased
D	10.04	9.98	10.02	9.97	10.04	Precise, unbiased

Graphical illustration The results of experiment represented by dot-plots. (The true value is 10.00).



Accuracy and Precision

- The FDA define precision as the closeness of agreement (degree of scatter) between a series of measurements obtained from multiple sampling of the same homogeneous sample under prescribed conditions.
- Barnhart et al [7] describes precision as being further subdivided as
  1. within-run, intra-batch precision or repeatability (which assesses precision during a single analytical run),
  2. between-run, inter-batch precision or repeatability (which measures precision over time)

Inter-Method Bias

- In the context of the agreement of two methods, there is also a tendency of one measurement method to consistently give results above or below the other method.
- Lack of agreement is a consequence of the existence of 'inter-method bias'.
- For two methods to be considered in good agreement, the inter-method bias should be in the region of zero.

Three Conditions

For two methods of measurement to be considered interchangeable, the following conditions must apply (Roy 2009):

- No significant inter-method bias
- No difference in the between-subject variabilities of the two methods
- No difference in the within-subject variabilities of the two methods (repeatability)

- To illustrate the characteristics of a typical method comparison study consider the data in Table I (Grubbs 1973).
- In each of twelve experimental trials a single round of ammunition was fired from a 155mm gun, and its velocity was measured simultaneously (and independently) by three chronographs devices, identified here by the labels '***Fotobalk***', '***Counter***' and '***Terma***'.

Round	Fotobalk [F]	Counter [C]	Terma [T]
1	793.8	794.6	793.2
2	793.1	793.9	793.3
3	792.4	793.2	792.6
4	794.0	794.0	793.8
5	791.4	792.2	791.6
6	792.4	793.1	791.6
7	791.7	792.4	791.6
8	792.3	792.8	792.4
9	789.6	790.2	788.5
10	794.4	795.0	794.7
11	790.9	791.6	791.3
12	793.5	793.8	793.5

Table : Velocity measurement from the three chronographs (Grubbs 1973).

- An important aspect of the these data is that all three methods of measurement are assumed to have an attended ***measurement error***, and the velocities reported in Table 1.1 can not be assumed to be ‘true values’ in any absolute sense.

Inter-Method Bias

- A simple estimation of the inter-method bias can be calculated using the differences of the paired measurements.
- The data in Table 1.2 (*next slide*) are a good example of possible inter-method bias; the ‘Fotobalk’ consistently recording smaller velocities than the ‘Counter’ method.
- Consequently one would conclude that there is lack of agreement between the two methods.

Round	Fotobalk (F)	Counter (C)	F-C
1	793.8	794.6	-0.8
2	793.1	793.9	-0.8
3	792.4	793.2	-0.8
4	794.0	794.0	0.0
5	791.4	792.2	-0.8
6	792.4	793.1	-0.7
7	791.7	792.4	-0.7
8	792.3	792.8	-0.5
9	789.6	790.2	-0.6
10	794.4	795.0	-0.6
11	790.9	791.6	-0.7
12	793.5	793.8	-0.3

Table : Difference between Fotobalk and Counter measurements.