Calculating the Standard Error of Measurement

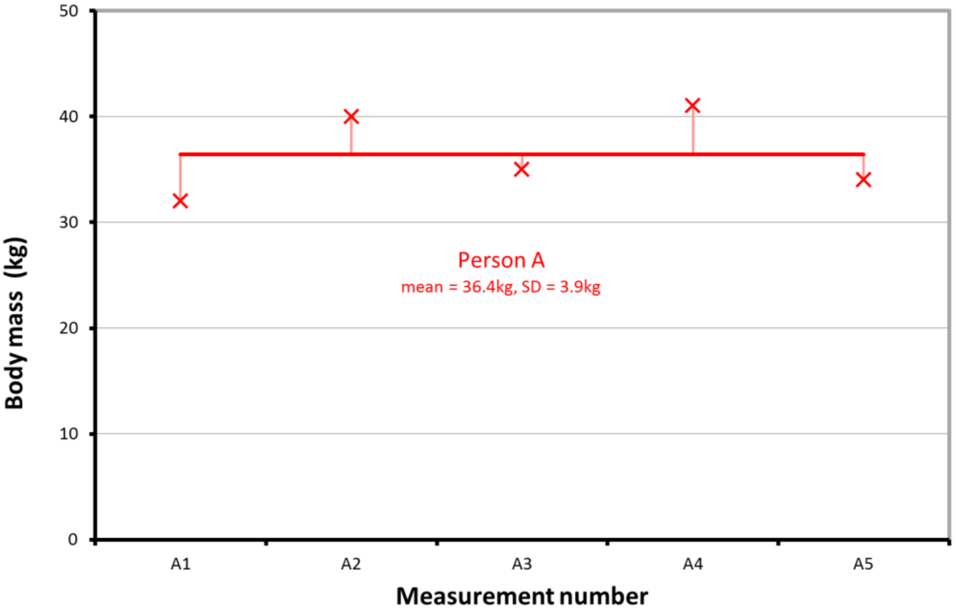
If we make a series of measurements of the same clinical variable we expect those measurements to show some variability. This may be because the measurement process is imperfect, or depend on exactly how the test is performed or because the underlying property being measured may vary from test to test. Understanding the amount of variability associated with a measurement can be very important when making clinical decisions. If we know that measurement variability is low then we can have confidence in any individual measure and act upon it. If the measurement variability is high then we will have less confidence. We might want to confirm the finding either by reference to different measurements or by repetition of the same measurement.

The simplest and often most appropriate measure of variability is the *standard error of measurement* (SEM). It is the standard deviation of a number of measurements made on the same person (indeed Bland and Altman prefer the term *within-subject standard deviation* [1]). Most textbooks suggest it is calculated as a derivative of the intra-class correlation coefficient (ICC) and as a consequence many people do not appreciate just how simple a measure the SEM is.

This short paper aims to demonstrates how to calculate the SEM directly from the data and it is hoped that this will help emphasise just how simple the measure is. It assumes that a balanced repeatability study has been performed. A balanced study is one in which the same number of measurements has been made on each person.

Calculating SEM from measurements on one person

The simplest possible calculation of the SEM is when we have a series of repeat measurements on just one person. In this example the body mass of one person has been measured on five occasions. In the graph below these are plotted as five crosses. The mean value (36.4kg) of those five crosses is plotted as a horizontal line. The five vertical lines represent the *deviation* of each measurement from the mean. The *standard deviation* (3.9kg) is a summary measure of how far all the measurements are from the mean. As measurements have only been made on one person the SEM is simply this standard deviation (3.9kg).



If we want to calculate this using a spreadsheet then the best thing to do is to arrange the measurements in columns and use the relevant spreadsheet functions to calculate the standard deviation.

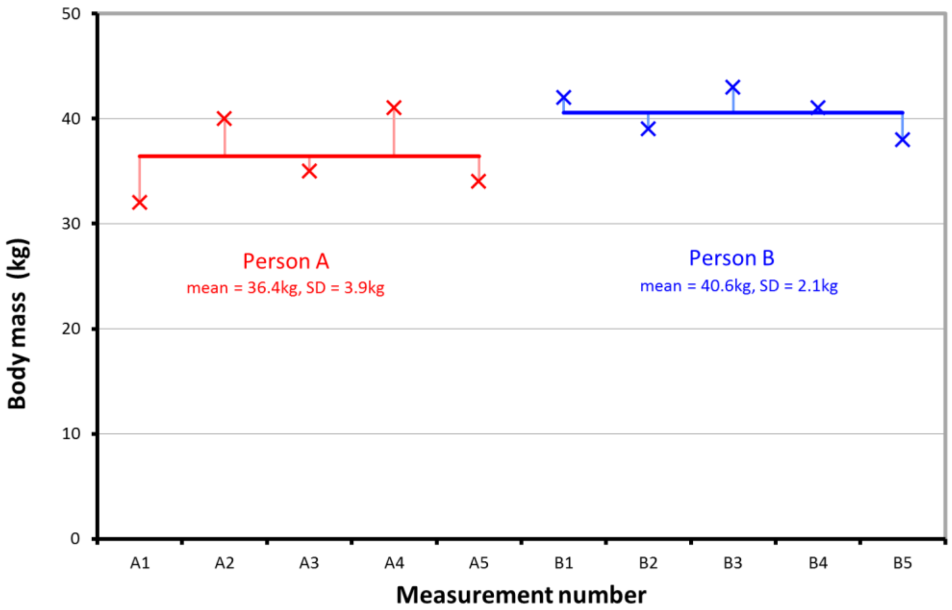
	Measurements						
	M1	M2	M3	M4	M5	mean	**SD**
Person A	32	40	35	41	34	36.4	**3.9**
Overall SEM							**3.9**

Calculating SEM from measurements on more than one person

Whilst calculating the SEM from just one person is quite possible it assumes that the variability in measurements is the same for everyone. It is therefore generally advisable to make measurements on a number of different people to test this. In this example we've thus also made measurements on a second person (in blue).

As might be expected, the mean values of the two people are different. You can see that the magnitude of the deviations (vertical lines) are also different indicating that the measurements are less variable for person B who has a smaller standard deviation (2.1kg).

Given that we now have two different within subject standard deviations we need to decide how to combine these to give us an overall SEM. It is tempting to take the simple mean of the two measurements but most statisticians agree that a better measure is the *root mean square* average. In this case this works out to be 3.1 kg ($= \sqrt{(3.9^2 + 2.1^2)/2}$).



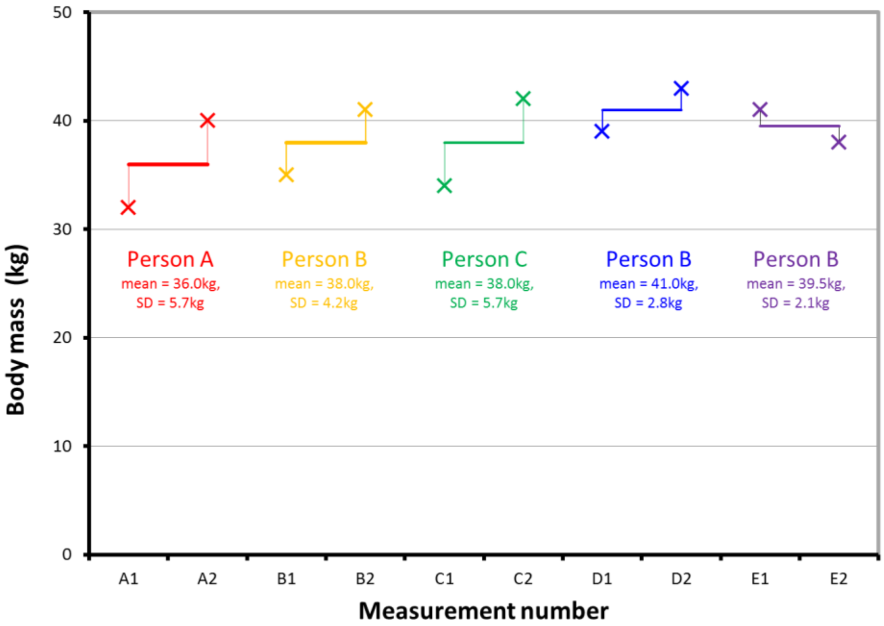
In a spreadsheet it is probably simplest to do this by adding an additional row.

	Measurements						
Measurement	M1	M2	M3	M4	M5	mean	**SD**
Person A	32	40	35	41	34	36.4	**3.9**
Person B	42	39	43	41	38	40.6	**2.1**
Overall SEM							**3.1**

You can adjust the number of measurements by adjusting the number of columns or the number of people by adjusting the number of rows but otherwise this format can be used to calculate the SEM from any balanced repeatability study.

Calculating SEM with just two measurements for each person

A very common design of repeatability study it to do just two measurements on each person. The results of such as study on five people are illustrated in the following figure.



To calculate the SEM we proceed in exactly the same way as we did for the previous study. Because there are only two measurements for each person part of this requires us to calculate the standard deviation of those two measurements. Although we would very rarely take the standard deviation of just two measurements in any other circumstances this is still valid and you should find that the corresponding spreadsheet functions still work.

	Measurements			
	M1	M2	Mean	**SD**
Person A	32	40	36.0	**5.7**
Person B	35	41	38.0	**4.2**
Person C	34	42	38.0	**5.7**
Person D	39	43	41.0	**2.8**
Person E	41	38	39.5	**2.1**
Overall SEM				**4.3**

As in the previous example the root mean square average is used to calculate the overall SEM ($=\sqrt{(5.7^2+4.2^2+5.7^2+2.8^2+2.1^2)/5}$).

Calculating SEM using Analysis of Variance (ANOVA)

I hope that the approach outlined above has persuaded you that the SEM is really simple concept, particularly if you think of it as the within-subject standard deviation. I hope you will also see that it is an extremely simple calculation to make if you format the data in a simple table.

You will, however, often see the SEM calculated using a technique called the *analysis of variance* or *ANOVA* for short. If you take the table of measurements from the 2 person data and perform an ANOVA you will obtain a table that looks something like this (depending on the spreadsheet or statistical package that you use).

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Person A	5	182	36.4	15.3
Person B	5	203	40.6	4.3

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	44.1	1	44.1	4.5	0.066688	5.317655
Within Groups	78.4	8	**9.8**			
Total	122.5	9				

You can ignore everything else on this table apart from the value of 9.8 highlighted in green. This is the within groups mean square (MS) value and the SEM is the square root of this (which equals 3.1 just as calculated in the section above).

If you regularly perform ANOVA this may be a quicker way of calculating the SEM than I've suggested above. If you have a really good understanding of ANOVA then this might help you understand what the SEM represents. If neither of these apply to you then it is generally easier just to calculate the SEM as suggested above. (It should be noted that using ANOVA to calculate SEM can be particularly cumbersome for the time series data commonly encountered in gait analysis).

A note for gait analysts

The outline I've given above will work for almost any data. (Technically the within-subject deviations have to be normally distributed but this is will be the case for most commonly encountered measurement data). Given that I'm a gait analyst however there are a couple of points that are worth adding. Most of these stem from the fact that most gait analysis data is supplied as a time series of data points across the gait cycle and that we often capture data from several cycles as part of any gait analysis session.

I'm often asked how the SEM should be calculated given all these factors:

- Should I chose a representative cycle or the mean of several cycles?
- How do I adapt this if my focus is on the peak power generation during the gait cycle rather than the value at a particular time?
- Should I calculate the SEM across the gait cycle or at one specific instant?

The simple answer is that you should process the data for a repeatability study in exactly the same way as you are going to process the data in your definitive study. Thus if when you analyse your data you take the mean of several cycles whenever you make measurements on an individual then you should do this when calculating the SEM. If your primary outcome measure is the peak of power generation wherever it occurs in the gait cycle, or is the impulse of a motion over stance then you should do this calculation first for the repeatability study and calculate the SEM of the measurement you are interested in.

There are a growing number of repeatability studies for biomechanical models which produce time varying outputs across the gait cycle. These tend to be providing data for generic use rather than in a specific context. It is my suggestion in this case that the SEM is calculated separately at each interval of the time normalised gait cycle (generally 50 or 100 data points). To do this it is often easier to re-format the data so that all the data from a particular time point can be represented on a single line. Data from the next time point can be input onto the next line and the analysis over multiple time points can be replicated very quickly. This can be further enhanced by putting data from other gait variables in the same columns one below the other in which case the whole analysis can be performed as the analysis of a number of long columns.

Person A		Person B		Person C		Person D		Person E		Standard deviations					**Overall SEM**
A1	A2	B1	B2	C1	C2	D1	D2	E1	E2	A	B	C	D	E	
32	40	35	41	34	42	39	43	41	38	5.7	4.2	5.7	2.8	2.1	**4.3**

The SEMs can be plotted across the gait cycle. For most kinematic variables the SEM is fairly constant and the average SEM across the gait cycle can be used as a broad indicator of the variability. For some other forms of data (particularly EMG) the SEM can vary quite substantially across the gait cycle and a single summary value may not be as useful.

References

1. Bland, J.M. and D.G. Altman, *Measurement Error.* British Medical Journal, 1996. 313:744-753.