Appendix 1 to Cadeddu M, Farrokhyar F, Levis C, et al. Users' guide to the surgical literature. Understanding confidence intervals. *Can J Surg* 2012;55(3)207-11.

DOI: 10.1503/cjs.010211

POPULATION VARIATION, STANDARD DEVIATION AND STANDARD ERROR OF THE MEAN

Standard deviation (SD) and standard error of the mean (SEM) measure 2 different things, but they are often confused. The following example can help explain the difference between them. If we draw a random sample and calculate mean and SD, the SD will tell us how spread out our observations are and how far the individual observations are from the sample mean. The larger the SD, the more spread out the observations are. If we draw several random samples from the same patient population and calculate the same statistic, for example a mean, it is highly unlikely that the calculated mean will be the same for each sample. Even if we control all biases, variations will occur purely by chance.[1] If we plot a histogram of means, they will form a sampling distribution of the mean. The SD of this sampling distribution is called SEM. The SEM simply explains, for a given sample size, how far the sample mean is likely to be from the true population mean. Therefore, it is the SEM that provides us with the information about an estimate of true population value and not SD. It is the SEM that is used for hypothesis testing and estimation.[1] The SEM can be computed from the SD by dividing the SD by the square root of the sample size used to compute the mean.

CALCULATION OF VAS MEAN DIFFERENCE FROM GROUP MEANS

We have learned from the body of the article (concept of confidence intervals [CIs]) that the estimation of population parameters (mean and variance) are often impossible and therefore are almost always unknown to us. That is why we conduct studies on a random sample of patients drawn from the whole population and assume that the results from these studies could be extrapolated as estimates of what might happen if the treatment were given to the entire population. For that reason, when we deal with a sample mean, we use a t-distribution to construct CIs because we could use the sample variance as an estimate of population variance in our calculation. The critical statistic-t for the α value of 0.05 (significance value) for a 2-sided test has different values for different degrees of freedom (df). When we have one variable, the df is sample size minus 1 (n – 1) because we lose 1 df for estimating the sample mean. For example, the df for 34 patients in the absorbable sutures group is 33 and for 29 patients in the nonabsorbable sutures it is 28. Looking at a table for t-distribution (this can be found in any statistical text book[2]), their corresponding statistic-t is 2.031 and 2.045, respectively. The formula for 95% CI for 1 sample mean is as follows:

$$(\text{mean} \pm t_{df}\,(\frac{SD}{\sqrt{n}})$$

Appendix 1 to Cadeddu M, Farrokhyar F, Levis C, et al. Users' guide to the surgical literature. Understanding confidence intervals. *Can J Surg* 2012;55(3)207-11.

DOI: 10.1503/cjs.010211

For this formula, t_{df} is the critical statistic-t for that specific df, *n* is sample size and SD is the standard deviation of the sample mean. Note that SD is the square root of the sample variance. The article has provided us with the mean VAS for cosmesis with their corresponding 95% CI for each group. As indicated in our discussion, means of both groups is not as useful in interpreting study results as the mean difference. In this case, determining the mean difference is useful, since this is the treatment effect, a value that is useful for the clinician. The calculation of the VAS mean difference is provided below.

For the absorbable sutures group, the mean and 95% CI was 79 (73–85) mm, and the number of patients included in this analysis was 34. For the nonabsorbable sutures, they were 66 (55–76) mm and 29 patients. We could use the previously mentioned formula to calculate the SD for both groups. Then, we could calculate the mean VAS cosmesis difference and its 95% CI between the groups. We are provided with enough information to calculate them. We need to work backward using the distance between the lower and upper boundaries of the CI to calculate SD for each mean.

Let's work this out for the absorbable sutures group. The CI around the mean is symmetric, and the distance between the mean and lower boundary is equal to the distance between the mean and higher boundary. For this group, this distance ($t_{df}(sd/\sqrt{n})$) is equal to 6 mm (79 – 73 = 6 mm or 85 – 79 = 6 mm). If we replace the value of 2.031 for critical statistic-t and 34 for *n* in this formula, the SD for this group would be 17.1 mm.

$$6 = 2.031 \times \frac{sd}{\sqrt{34}} \rightarrow \qquad sd = \frac{6 \times \sqrt{34}}{2.031} = 17.1$$

If we repeat the same process for the nonabsorbable sutures group, we would find that the SD for is equal to 29.4 mm, which is much larger than SD 17.1 mm for the absorbable sutures group. Now that we have both means and SDs for each group, we could calculate the mean difference and its corresponding 95% CI. The formula to construct 95% CI for a mean difference is as follows:

$$(m_1 - m_2) \pm (t_{df}) \sqrt{\frac{sd_1^2}{n_1} + \frac{sd_2^2}{n_2}}$$

To calculate the mean difference and its 95% CI from the above formula, we only need to find the appropriate value for t_{df}. When we have 2 groups and their variance is equal, the df would be equal to $[(n_1 - 1) + (n_2 - 1)]$. However, the variance in the absorbable sutures group ($17.1^2 = 291.8$) and that in the nonabsorbable sutures group ($29.4^2 = 848.1$) are far from being equal ($p = 0.002$, so the hypothesis of equality is rejected); therefore, we have to use the following formula to calculate the approximate df:

Appendix 1 to Cadeddu M, Farrokhyar F, Levis C, et al. Users' guide to the surgical literature. Understanding confidence intervals. *Can J Surg* 2012;55(3)207-11.

DOI: 10.1503/cjs.010211

$$df = \frac{[(\frac{sd_1^2}{n_1}) + (\frac{sd_2^2}{n_2})]^2}{[((\frac{sd_1^2}{n_1})^2 /(n_1 - 1)) + ((\frac{sd_2^2}{n_2})^2 /(n_2 - 1))]} = 43.66$$

The statistic-t for df of 43.66 is equal to 2.015. If we incorporate the right numbers in the formula, we find that the 95%CI for a mean difference of 13 mm is 0.6–25.4 mm.

$$(79 - 66) \pm (2.015)\sqrt{\frac{(17.1)^2}{34} + \frac{(29.4)^2}{29}} = 13 \pm 12.4 = (0.6,\ 25.4)$$

Therefore, the mean difference of the VAS score is 13 mm with a 95% CI of 0.6–25.4 mm.

Please note that, most often, authors provide the mean difference, risk difference or odd ratios with the corresponding CIs, and there is no need for further calculations.

References

1. Akobeng AK. Confidence intervals and p-values in clinical decision making. *Acta Paediatr* 2008;97:1004-7.
2. Pagano M, Gauvreau K. Chapter 9. Confidence intervals. In: *Principles of biostatistics*. 2nd ed. Florence (KY): Brooks/Cole Duxbury; 2000. p. 241-231.