

**AN INVESTIGATION INTO THE RELATIONSHIP BETWEEN HEMORRHOID
DISEASE AND THE USE OF SMARTPHONES IN THE LAVATORY**

Unique Protocol ID: 12/22/2017 – 051

Date: February 11, 2018

Statistical Analysis

All data will be analyzed using the SPSS package program (version 20). The data will consist of categorical, nominal, and continuous numerical variables. Frequency and percentage rates will be given for categorical and nominal variables. Following testing for normal distribution, variables with normal distribution will be expressed as mean \pm standard deviation, while those that do not exhibit normal distribution will be expressed as median and minimum – maximum values. The relationship between risk factor and outcome variable will be evaluated using the Chi-square test. The effect of smartphone use on the presence or absence of hemorrhoids will be analyzed using the Binary Logistic Regression method. Any difference in the degree of hemorrhoids with respect to the duration of exposure to the risk factors will be determined by ANOVA or the Kruskal-Wallis test. A *p*-value of less than .05 will be considered statistically significant.

Calculation of Sample Size

According to the results of the pilot study, 36.4% of the population with hemorrhoidal disease use a smartphone for at least 2 minutes while in the lavatory. In the control group, this percentage was only 13.2%. Since this is a case-control study, the investigators first calculated the number of cases required to calculate the prevalence of a group with hemorrhoidal disease as per the below:

$$n = \frac{Z^2 p(1-p)}{d^2} \quad n = \frac{1.96^2 * 0.364(1-0.364)}{0.05^2} \quad N = 355.7$$

According to the above, in order to estimate a 5% difference in effect size, with an error rate of 5% and 80% power, at least 356 patients are needed. However, in this study, the sample for the logistic regression analysis needs to be calculated in order to investigate to what extent the use of a smartphone during defecation increases the risk of hemorrhoids.

According to analyses performed using the pilot study data, the percentage of the whole population who use a smartphone for at least 2 minutes while in the lavatory is 26.9%. In the logistic regression analysis, in order to determine whether or not smartphone use during defecation (the independent variable) has an effect on the presence or absence of hemorrhoidal disease (the dependent variable), the approach suggested by Peduzzi et al. (1) is as per the below:

$$N = 10 k / p$$

where

P = the rate of positive or negative events in the population

K = the number of independent variables.

$$N = 10 * (22 / 0.269) = 817.8$$

According to these calculations, calculating the prevalence of smartphone use in the lavatory among the hemorrhoid patient population and measuring the effect of this habit on hemorrhoid development requires 818 patients and the same number of controls. Taking into account the possibility for incomplete or incorrect questionnaires by adding 10% loss to this figure results in a total of $818 + 81 = 899$ participants needed.

1. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology* 49: 1373-1379.