**Academic Exchange project in Russia Report**

First, I would like to say thanks to Nagasaki University and Prof. Takamura for giving me this opportunity to Academic Exchange project in Russia.

1. **Introduction**

Date : 25 January- 1 February 2019

Place: North Western State Medical University, St. Petersburg, Russia.

We have 8 students from Nagasaki University and 2 students from Fukushima

1. **Purpose**

During I have learned and read the all presentations of the Bio-Statistics course, it was very interesting and curious to know what are statistics. Is important to know there are two kinds of variable that we can work with. One of them is Categorical data and the other is Quantitative.

Quantitative data are all that we can measure and represent using numbers. To analyze them we have two important features, the central values and the spread about these values. The measures of central values are mean, median and mode. These data can give us information about the distribution of the data we are working with (Normal distribution or non-normal distribution). Measure of variation are Standard deviation and Range (inter-quartile range). Commonly we can represent the data by a histogram or boxplot.

In the other hand, Categorical data can be analyses using frequency tables and bar chart.

Statistic can be use in epidemiological approach when we are interested in making statement about a population but most of the case the resources are not enough to study the whole population so, we need to work with samples and infer about the nature of the population from these samples. Sample have to represent the population. Is important trying to reduce the bias using random sampling method. Therefore, we need to work with confidence intervals in most of the case to have a biggest probability to contain the mean of the true population (usually 95%).

We use the statistics data to make the testing of scientific hypothesis, to judge whether a hypothesis is reasonable or not we can evaluate probability of the observation data if the hypothesis was true. This probability is called a P-value.

Nowadays there are a lot of software we can use to determinate the statistical values and in case we have a hypothesis, we have difference test that we can easily use to calculate the p-value. Some of them like R-studio or Stata are powerful to work with a big amount of data and friendly with the operator in case the person does not have informatics background.

The choice of measure depends on the nature of the data and the purpose of the analysis. For instance, the arithmetic mean is useful for statistical inference but is easily influenced by outliers. Median is not affected by outliers and is useful for skewed data.

**Conclusion**

Statistic are very useful in all the fields and for me is very important because using the statistical knowledge I will be able to create database in Cambodia with variables related to radiation issues. So it can be helpful not only in my country as a government office, as well in the University in the department of Energy Technique and Nuclear Science. We can make research that will provide important information to all the countries and try to improve the actual situation.

**Name: Thon Mao,**