Student: Nessipkhan Arman

I learned many new methods during this course. For me as a young specialist, it was very important to listen such an experienced professional and a wise person. In my opinion the main points were:

1) Modern statistical methods
2) Methods of quantitative and qualitative analysis
3) Different type of statistical samples
4) Statistical software: programming R and SAS
5) Logistic regression
6) Basics of statistics in the research

After listening to this course, I have a clear image of what biostatistics is. The lectures were engrossing, I enjoyed the combination of presentations, samples and cases.

I think for many who decide to connect their fate with the science of young researchers like me, questions arise in my head:
What is statistics?
Biostatistics?
How can I apply this in my research?

As the professor said, without an adequate understanding of the basics of statistical analysis, neither planning nor processing the results of a scientific experiment is possible. And I absolutely agree with that. Since all scientific work is based on a clear understanding and planning of our research. For a long time studying at a medical university, I have repeatedly come across statistics and, although in theory, understood what statistics and its role are, but did not understand how to apply it in my research.

History is the basis of everything - the first day began with the basic concepts of statistics and its foundation. I was very lucky to meet a professional in my field and an impeccable lecturer. We had a week of productive work, because after the lecture we had very interesting tasks that underlined the topic of the lecture. It was a very good idea to complete the task after each lecture - it's so quickly absorbed.
Every day after class, I started to work, I especially liked working with the R Studio program since it is easy to use. It's a pleasure to work on it like working with a calculator.
In my opinion, I have opened up great opportunities for a better analysis of my research project, and I realized that I need to measure all my indicators and degrees of variability. I already applied Student Test criteria in my research, although in the process of research I was also able to test the null hypothesis.

Consequently I am starting to analyzing of quantitative and qualitative variables in my research. All things considered in the lecture its very important for my research.
The field of statistics is the main part of the work of any researcher. I have learned programming software R and SAS.
For me personally, statistics is of great importance in my work. Including planning, collecting materials, analyzing and processing research data, even until the conclusion I will apply statistical methods and programs.

The most important question is how can this be learned?
I'll name the answer: "7000 km or 7 days of highly effective statistics."
I want to thank you for your great contribution and opportunity to the future of my scientific life!

Imakhanova Aiganym

How often do we read an article in a scientific journal or news on the Internet and hear the phrase "According to statistics ..."?

Biostatistics, being a relatively young science, has taken a rather solid place in our lives. As a doctor, I have repeatedly quoted data from the sites world-statistics.org, egov.kz, worldofoncology.com and others. However, it was not always clear to me what exactly these numbers give and what they mean in clinical practice. Our weekly course of biostatistics from North-West State Medical University. I.I. Mechnikov helped me to finally understand the issues of biostatistics and its definitions, and even improve some practical skills.

Our teacher, Professor S. L. Plavinsky, began our course with the basics of biostatistics and the main terms. So, our first lecture was devoted to the history of biostatistics and to some scientists whose discoveries made a significant contribution to the development of biostatistics as a science. We also learned how scientific work is done. First of all, we should create a research plan. The program includes the topic, purpose and objectives of the study, formulated hypotheses, determination of the object of study, units and volumes of observations, a description of statistical methods for forming a sample, collection, storage, processing and analysis of data, research methodology, as well as a list of used statistical tools. Hypotheses (expected results) are the most important component of the program. Hypotheses are formulated using specific statistical indicators. The main requirement for hypotheses is the possibility of testing them in the research process. The results of the study can confirm, correct or refute the hypotheses put forward.

Subsequent lectures were devoted to the main definitions and methods used in data processing. As a graduate student, I know that statistical data analysis is an important part of any scientific research. It must be used both at the stages of work planning and directly in the analysis of data. The use of various methods of statistical analysis of individual indicators obtained during the study significantly increases the value of available information and facilitates its perception: data can be structured, visualized - displayed in the form of tables and graphs. For statistical analysis of biomedical data, they are used as standard commercial statistical packages (for example, SAS, SPSS, STATISTICA, MedCalc, etc.) and free software (for example, R package, Hierarchical Clustering Explorer, Dendroscpe, etc.) For me individually, the most interesting lecture was statistical software: the introduction in R and SAS. During this lecture I learned how to work in the R program. The first version of R was released in 1995. In contradistinction to SAS, which is based on a procedural statistical language, R is an object-oriented programming language. This makes R program very extensible. Due to the relative ease of installation and free and open access, the ability to find the answer in a difficult situation on the Internet and the convenience of work, I think that I will use R in my future research work.

In summary, I think that the biostatistics course was very informative for me. The knowledge of medical statistics is necessary for understanding the biomedical processes and phenomena, the laws of their manifestation, understanding the logic underlying the diagnosis, treatment and prognosis of various diseases. Statistical analysis allows one to substantiate one or another tactic of a doctor in the prevention or treatment of diseases. In addition, a huge flow of information can be easily represented as table, graph etc.

Kazakova Darya

Biostatistics report

The purpose of my research work is to measure the autophagy activity on cancer cells after exposure to the various drugs and radiation. To do this, I use immunohistochemistry with subsequent analysis under a fluorescence microscope, where I can collect data using the built-in software which calculates the amount of contrast. Then I upload the image of cells from each sample and read the information. As a result, I get dependent continuous quantitative variables.

For the statistical analysis I use Excel with a ToolPak add-in program that calculates a huge variety of parameters, such as mean, median, mode, variance, standard deviation. The set of statistical parameters can be estimated in “descriptive statistics” mode. Next, I can plot histogram that shows the distribution law. In Excel I also can compare the samples using averages or variances: t-test, F-test.

Speaking about Excel statistics, we think about normally distributed (or approximately so) - “parametric” tests. Normal distribution is essential for these tests to be valid. It is symmetric bell-shaped frequency distribution, in which mean, median, and mode all have the same value. If the data do not conform to a normal distribution, “nonparametric” tests should be used. To perform nonparametric tests in order to do all calculations fast and exclude manual entry of formulas it is better to use special programmes such as Microsoft R and SAS. In these programmes I just need to call a function and pass the data as a parameter. For example, `shapiro.test(data)` function used for Shapiro-Wilk test for checking the normality of distribution, `wilcox.test(group1, group2)` for Wilcoxon test to compare two paired groups.

Since this is my first serious scientific work and I have poor experience in statistical data analyses, I would like to master my knowledge. After the course of Biostatistics, I finally understood the basic statistical concepts and formulas that I tried to understand during my previous study. I would like to go deeper into exploring Microsoft R and SAS to apply this knowledge in my future research work.

Division of Disaster and Radiation Medical Sciences
Nagasaki University Graduate School of Biomedical Sciences
Junichi Yoshida

Statistics in clinical studies of radiological imaging

I go to graduate school as a member of society. I usually work as a radiologist at a hospital. My clinical research focuses on reducing lens exposure in Computed Tomography (referred to as CT) examination. CT examination in Japan are increasing year by year, and medical exposure is a problem. In particular, lens exposure has been reduced due to the 2011 ICRP recommendation, which has attracted attention. My research examined whether the Imaging Plate (referred to as IP) could be used as radiation protection. It is important that this study measures the exposure dose and that the reduced dose images are secured for clinical use. The image quality evaluation includes a physical evaluation and a visual evaluation. The physical evaluation uses a quantitative evaluation, while the visual evaluation is a qualitative evaluation and requires a statistical method.

In this visual evaluation, 7 groups of 7 evaluators evaluated how much low-contrast substances can be recognized in the captured images. Evaluate how many objects up to 8 can be seen. The statistical method used the multiple comparison test (Steel-Dwass).

A	B	C	D	E	F	G
7	6	7	5	5	5	6
7	7	8	7	8	8	6
7	6	7	7	7	6	6
7	6	8	7	7	6	6
7	7	8	7	8	7	6
7	6	7	7	7	7	6
7	6	8	7	8	6	7

Steel-Dwass

	n	rank sum	rank average
A	7	206.5	29.5
B	7	111.5	15.92857143
C	7	270.5	38.64285714
D	7	179	25.57142857
E	7	227	32.42857143
F	7	138	19.71428571
G	7	92.5	13.21428571

Test result

	rank sum	pair n	expectation	variance	rank sum − variance	level of significance 5% critical region	level of significance 1% critical region	test statistic	5% reference point	1% reference point	
A,B	70	14	52.5	42.40385	17.5	19.19869041	22.47989602	2.687419249	2.948282	3.452166	
A,C	38.5	14	52.5	37.69231	−14	18.10069891	21.19424922	−2.28035085	2.948282	3.452166	
A,D	56	14	52.5	12.25	3.5	10.31898603	12.08258106	1	2.948282	3.452166	
A,E	45.5	14	52.5	38.5	−7	18.29360701	21.42012682	−1.12815215	2.948282	3.452166	
A,F	63	14	52.5	44.55769	10.5	19.68023706	23.04374274	1.572997216	2.948282	3.452166	
A,G	73.5	14	52.5	45.23077	21	19.82832219	23.21713678	3.122498999	2.948282	3.452166	*
B,C	31	14	52.5	54.51923	−21.5	21.76926871	25.48980616	−2.911813801	2.948282	3.452166	
B,D	41	14	52.5	47.25	−11.5	20.26608626	23.72971812	−1.673003825	2.948282	3.452166	
B,E	38	14	52.5	55.32692	−14.5	21.9299297	25.67792536	−1.94939453	2.948282	3.452166	
B,F	50.5	14	52.5	48.59615	−2	20.55274952	24.06537437	−0.286899008	2.948282	3.452166	
B,G	56	14	52.5	31.09615	3.5	16.44077624	19.25063285	0.627645914	2.948282	3.452166	
C,D	68	14	52.5	43.75	15.5	19.50105059	22.83393191	2.343379733	2.948282	3.452166	
C,E	57.5	14	52.5	49	5	20.63797207	24.16516212	0.714285714	2.948282	3.452166	
C,F	69	14	52.5	55.32692	16.5	21.9299297	25.67792536	2.218276534	2.948282	3.452166	
C,G	75.5	14	52.5	53.84615	23	21.63447318	25.33197302	3.134371662	2.948282	3.452166	*
D,E	43.5	14	52.5	44.42308	−9	19.65048612	23.00890713	−1.350324636	2.948282	3.452166	
D,F	58.5	14	52.5	49.26923	6	20.69459213	24.23145898	0.854797729	2.948282	3.452166	
D,G	67	14	52.5	49	14.5	20.63797207	24.16516212	2.071428571	2.948282	3.452166	
E,F	63	14	52.5	56.53846	10.5	22.16873815	25.95754804	1.396424004	2.948282	3.452166	
E,G	68.5	14	52.5	54.65385	16	21.79612778	25.52125565	2.164260921	2.948282	3.452166	
F,G	57	14	52.5	44.55769	4.5	19.68023706	23.04374274	0.674141664	2.948282	3.452166	

* 5% significant

** 1% significant

As a result, a significant difference was recognized at a level of significance 5% between A/G and C/G.

DISPATCH TO ST. PETERSBURG

Kot Palina, Disaster and Radiation Medicine Master's Course, Nagasaki University

From January 22 to February 2, 2020, we went to St. Petersburg to participate in a seminar on biostatistics. It was a course consisting of 14 lectures, and covering the latest and current trends in the use of statistical methods in biomedical research. The seminar was held at the North West State Medical University (NWSMU) named after M.Mechnikov.

Departure from Nagasaki took place on January 22, and we arrived in St. Petersburg in the evening of January 23. On the same day, representatives of NWSMU met us at Pulkovo Airport and helped to establish cellular communications and the Internet in our mobile phones. After that, we were taken to the NWSMU Academy Hotel.

On January 24, after an organized breakfast at the hotel, representatives of the NWSMU international department, together with employees of the NWSMU History Museum, conducted a tour of the Medical University Museum. After that, in the assembly hall of the main building, the NWSMU rector said a welcoming speech, and Professor Plavinsky gave lectures on the history of biostatistics and probability theory. Then a tour of the city of St. Petersburg was organized with sightseeing in the city centre. We visited Nevsky Prospect, a monument to Alexander Nevsky, a monument to Pushkin, the Main Theatre of St. Petersburg, the Neva River, and central bridges.

On January 26, I visited the Hermitage with a tour of the expositions of English, German, Danish, Egyptian, Russian, Italian, Spanish works. The museum is very large, it presented paintings and sculptures, as well as dishes, murals and other forms of art. Even the building of the museum inside and out was distinguished by sophistication of architecture.

On January 25 and 27-31 there were lectures about Observation and Experimental design in health sciences; Statistical software, R and SAS programs; Descriptive statistics; Inferential statistics; Hypothesis testing and confidence intervals; Z, t-tests; Analysis of Variance; Linear regression; Table analysis; Chi-square, Fisher's exact test; Logistic regression and long-linear models; Survival analysis; Robust statistics, Complex survey design, Qualitative research methods, Text mining.

After each lecture, there was a discussion and homework. After the whole course, there was a graded test.

All lecture materials and R and SAS code writing software were provided. At the end of the seminar, all participants received a certificate of successful completion of the course on biostatistics.

Despite of exhausting flight this trip was worth it, because I gained valuable knowledge on statistics, received answers to questions that worried me (some of them have worried me for years), to which I did not receive answers even during the Research Methods Advanced course last year. Professor Plavinsky answered for all my questions, and spent a lot of time to explain me practical points of choosing statistical methods after compulsory lectures. The discussion about using statistics in our research was the most useful, but, unfortunately, there was not enough time to learn all the methods. In addition, we felt a lack of time to gain practical skills in using the discussed software.

"Biostatistics Report"

North-Western State Medical University named after I.I.Mechnikov, Saint-Petersburg, Russia, 24-31 January 2020

12 February 2020

Name: THU ZAR WIN

University: Nagasaki University

When I and my friends were landed to St. Petersburg, I can't believe with the night light of city. The way we go to Hotel from Pulkovo Airport (Saint Petersburg), everywhere was very bright and amazing architecture building. The night when we arrived the city, there were first time snowing during this winter in Saint Petersburg. And also there is one thing which I have never seen like that in any other countries is that so many fume and smoke inside the city. It was very fascinating and amazing Biostatistics seminar which were held in St. Petersburg, Russia. Besides the Biostatistics seminar, I did some questionnaire to Master/ PhD military students from Myanmar who are studying astronautics in Moscow about health care system in Russia. I asked to them so many questions which is related to Russia's Health system and Cultures. During this trip, I learned not only about the Biostatistics but also the Russia's Health system and Russian people mentally.

Seminar in St. Petersburg was very fruitful and grateful which is my first time learning about Biostatistics. Professor Plavinsky answered all the following questions what I have expected to know before I visit to North-West Medical University named after I.I Mechnikov, Saint-Petersburg, Russia: "Why I need to know about Biostatistics/ statistics? How would be able to apply the knowledge in my research which I have received from your seminar?". During the second day of his seminar, it was totally stressful lecture for me because I had no clue about statistics, statistical tests and statistical significance. However, in the next days of his seminar, I have seen the point that statistics are helping us to understand the world and the people's psychological and sociological. By understanding the statistics, we (researcher) can be informed, evaluate the credibility and the usefulness of information, and make appropriate decisions. In scientific studies, we can use the statistic for data collection, analysis, interpretation, explanation and presentation. By using statistics, it will guide us in our research for proper characterization, summarization, presentation and interpretation of the result of research. It can provide a platform for research: How to go about my research, either to consider a sample or the whole population, the Techniques to use in data collection and observation, how to go about the data description (using measure of central tendency). Finally, the main point what I concluded from your seminar is that statistics can help us to understand and describe phenomena in our word and help us to draw conclusions about those phenomena. Also on the other side of your seminar, I noticed the following misuses of statistics in a research: biased samples, proof of null hypothesis, misleading graphs, data manipulation, misreporting or misunderstood of estimated error, overgeneralization, loaded questions, discarding unfavorable data and false causality.

Currently I am doing my research which is mainly focus on the "**Risk Communication in the Recovery Phase after a Nuclear accident: the contribution of the Co-expertise process**

model". The objective of my research is better identify the factors and mechanisms that help restore confidence among the people affected by the accident as well as their anatomy. Despite I will not use the concrete statistics tools in my research, it is important to understand key aspects of Biostatistics and the result of epidemiological studies, especially in the fields of radiation where there are uncertainty and there are still lots of debates of the low level of exposure. In the first day of his seminar, as he mentioned in the first motivating example which is about carcinogenic risk induced by low doses of ionizing radiation is controversial and it cannot be assessed with epidemiologic methods alone because at low doses the data are imprecise. As radiation helps in diagnosing and treating disease, but also causes diseases, it is important to understand and use radiation wisely and important to keep track with the results of epidemiological studies and it is very important to set the policies. By using the key statistics, I may be doing questionnaire on the affected people/ stakeholders and interrogating people socially which is generally done by the use of questionnaire or be the use of enquiries. I can do statistics with that and looking at average and standard deviation with the concept of statistics. These are all what I can do with statistics in my research. As Prof.Plavinskin urge me to see some literature on using qualitative methods in studying my questions, I may study the people perceptions of some risks and also I may use Likert-like scale to have (semi-) quantitative data in my research. I do greatly appreciate him for his amazing Biostatistics seminar in Saint-Petersburg, Russia and also for his advice on my research. The visit was the amazing opportunity for me because I have learned not only Biostatistics but also Russia's Health Care System, Russian people mentally, Russia's Histories. I will never forget beautiful memories what I have experienced in Saint Petersburg.

Sailaubekova Yerkezhan

«Using statistics in research work»

In October, I started my research at the Nagasaki University. I am work in the Department of pathology and study violation of the thyroid gland. As you know, scientific work includes a theoretical and practical part. During the discussion of the theoretical part, we often refer to foreign articles to determine the relevance of this topic. In the results and discussions section often provided: tables, graphs, and charts , which at that time seemed complicated to me. I realized that after completing the practical part, which consists of DNA extraction, DNA sequencing and immunofluorescence analysis, I would have to process the data and demonstrate it in the form of graphs, etc.To my great happiness, we just had a week-long course in biostatistics at the North West State Medical University named after I.I Mechnikov. Having successfully completed this cycle, I can confidently say that I will apply this knowledge in the presentation of my scientific work.

So what is biostatistics? Biostatistics – it is statistics applied to the living world. It includes demography, epidemiology and clinical trials. Another name-biometrics. In simple terms, it is everywhere in real life. In my research work, I will form a null hypothesis and prove it by using the next methods : Paired t-test, Wilcoxon test, descriptive statistics, Cohran-Cox t-test. In cases of comparison of more than 2 groups I will use ANOVA. Analysis of variance it is method of man's comparison though variances are analyzed. Main idea is to calculate variance without taking into consideration group membership all from one population and compare with sum of intergroup variances. If groups from different populations then variance ignoring group membership will be much higher than sum of intergroup variances. All the methods which I mentioned above I will be processed in the Microsoft R program. Because, R is free and open access, supported by a very large and active worldwide user community. Participate in online forums to address questions from fellow R-users highly adaptable and customizable. This facilitates keeping a record of the precise steps and methodologies used in my analyses, making work reproducible. Also, in the future, I plan to analyze data in the SAS program, which is one of the most popular products for collecting, analyzing and providing data. The system itself consists of several modules that perform certain tasks. The most commonly used modules are the BASE module, which provides interaction with the user and without which the system does not work and the STAT module, which includes various statistical programs.

In conclusion, I would like to say that I will apply my knowledge on biostatistics directly in my research activities.