

## Анализ данных системы автоматического круглосуточного мониторинга PhenoMaster с помощью U–критерия Манна–Уитни

И. А. Бухтеева\*

*Московский государственный университет имени М. В. Ломоносова,  
физический факультет, кафедра общей физики и молекулярной электроники  
Россия, 119991, Москва, Ленинские горы, д. 1, стр. 2*

(Статья поступила 03.05.2018; Подписана в печать 21.06.2018)

PhenoMaster — это уникальная модульная автономная платформа. Она предоставляет ученым разных исследовательских областей современную технологию для полностью автоматизированного и синхронизированного метаболического, поведенческого и физиологического мониторинга — косвенной газовой калориметрии, оценки потребления пищи и жидкости, локомоторной активности, обучения, памяти и т. д. В результате измерений, проводимых системой PhenoMaster, имеется большой объем экспериментальных данных, который тяжело обработать и проанализировать с помощью персонального компьютера.

В данной работе был предложен метод анализа измерений на основе U–критерия Манна–Уитни, который был применен для обработки комплексного исследования влияния перорального введения биологически-активной добавки на основе наночастиц серебра на физиологические функции млекопитающих. Физиологические показатели животных фиксировались в системе автоматического круглосуточного мониторинга PhenoMaster. Полученные данные исследовались на наличие изменений в количестве потребляемых воды и пищи, изменений массы тела, нарушений циркадных ритмов, изменений в двигательной активности. Предложенный метод учитывает погрешности измерений, особенности живых организмов, такие как суточные ритмы, а также использует усреднения по времени и по ансамблю, что делает возможной обработку больших объемов полученных данных на персональном компьютере и приводит к релевантным результатам.

PACS: 87.50.sg, 07.05.Kf, 29.85.+c

УДК: 51-76, 519.25, 519.23, 51-73, 53.087.45

Ключевые слова: PhenoMaster, анализ экспериментальных данных, обработка данных, статистические критерии, U–критерий Манна–Уитни, непараметрические критерии.

### ВВЕДЕНИЕ

В настоящее время в биологии и медицине при проведении экспериментов используется современное оборудование не только для получения точных результатов, но и для автоматизированного мониторинга сразу нескольких параметров. Такое современное оборудование позволяет сократить материальные и временные расходы на проведение эксперимента и увеличить достоверность его результатов. В связи с этим, в настоящее время для проведения экспериментов на живых объектах, связанных с изучением двигательной активности, потреблением пищи, обучением, изменением веса и др. стали чаще использоваться модульные автономные платформы, такие как IntelliCage, PhenoMaster [1] и др. Они предоставляют современную технологию для полностью автоматизированного и синхронизированного метаболического, поведенческого и физиологического мониторинга. Запись данных может производиться за любой заданный временной промежуток, например каждую секунду, и в течение любого заданного периода [2–4].

Измерение нескольких физиологических величин одних и тех же животных позволяет определить их корреляцию и исключает влияние индивидуальных пара-

метров каждого из организмов на нее. По итогам проведенного с использованием таких платформ эксперимента имеется большой объем данных, который затруднительно обработать и проанализировать на персональном компьютере. Кроме того, существует проблема учета индивидуальных характеристик исследуемых животных, таких как суточные ритмы, внезапная болезнь, особенности поведения. Живые объекты тяжело описывать математически, так как это требует введения большого количества зависимых друг от друга параметров. Более того, анализ данных, полученных с помощью системы PhenoMaster, выходит за критерии применимости широко распространенных статистических методов обработки, например таких, как t–критерий Стьюдента. Программного обеспечения для анализа данных, разработанного специально под модульные платформы, на момент выполнения данной работы не существует. Подобные проблемы ранее решались с помощью написания отдельных программ на языке R. То есть для обработки каждого конкретного исследования требуется создание новой программы, что занимает много времени и неэффективно при проведении серии различных экспериментов [5, 6].

В данной работе был предложен способ анализа данных на основе U–критерия Манна–Уитни для независимых выборок, продемонстрированный на биологическом эксперименте с использованием платформы PhenoMaster [7].

\*E-mail: [iabukhteeva@mail.ru](mailto:iabukhteeva@mail.ru)

## 1. ОПИСАНИЕ ЭКСПЕРИМЕНТА

Анализ данных проводился на основе комплексного исследования влияния одно-, двух-, четырех- и шестимесячного ежедневного перорального введения биологически-активной добавки на основе наночастиц серебра марки «Арговит-С» на поведенческие, когнитивные и физиологические функции млекопитающих [8]. Препарат представлял собой водную дисперсию наночастиц серебра с концентрацией 10 мг/мл, стабилизированных полимером медицинского назначения поливинилпирролидоном. В качестве модели млекопитающих в работе использовали мышей-самцов линии С57В1/6 в возрасте 2 месяцев, массой 19–27 г на начало эксперимента. Животных разделили на четыре экспериментальные группы: «30 дней», «60 дней», «120 дней» и «180 дней» по 20 мышей в каждой группе, состоящих из двух подгрупп: «Контроль» и «Серебро» (по 10 мышей в каждой). Мыши групп «Серебро 30», «Серебро 60», «Серебро 120» и «Серебро 180» получали наночастицы перорально в виде водного раствора в количестве 50 мкг/сут в течение соответствующих периодов времени. Мыши групп «Контроль» наночастиц серебра не получали.

В ходе эксперимента физиологические параметры животных (потребление пищи и жидкости, локомоторная активность, вес) фиксировались ежесекундно в течение первого дня адаптации к клетке и экспериментального времени равного неделе. Измерения проводились в системе автоматического круглосуточного мониторинга PhenoMaster, которая была выбрана как наиболее подходящая для контроля поведения животных, по сравнению с IntelliCage, которая в основном используется для фиксации психологических и когнитивных параметров живых организмов.

После получения данных были сформулированы следующие задачи обработки и анализа:

1. Отсутствие универсальных статистических методов обработки для живых объектов.
2. Большой объем полученных данных.
3. Необходимость учета биологических циклов: потребление пищи и воды, сон и активность животных является циклической – то есть проявляется в строго определенных промежутки времени как максимумы и минимумы.
4. Необходимость учета индивидуальных характеристик каждого животного.
5. Влияние возраста животных на результаты.

При проведении эксперимента каждый физиологический параметр непрерывно фиксировался каждую секунду на протяжении недели, что привело к большому объему итоговых данных. Заметим, что полученные данные нельзя обрабатывать стандартными статистическими методами такими как  $t$ -критерий Стьюдента,

который требует нормального распределения и равных дисперсий [9, 10]. Для точного ответа на вопрос оказывают ли наночастицы серебра влияние на животных требуется сравнивать не только средние активности и потребление, но и циклы животных, а именно нет ли смещения или изменения продолжительности циклов жизнедеятельности, не появились ли новые пики активности [11, 12]. Также необходимо учитывать индивидуальные характеристики животных, например некоторые мыши активнее остальных, другие потребляют больше воды при ответе на вопрос о влиянии приема наночастиц на живой организм. При анализе данных эксперимента нельзя забывать и о возрасте объектов — для мышей срок в 120 и 180 дней является существенным по сравнению с продолжительностью жизни, за это время в их организме происходят ощутимые возрастные изменения, которые необходимо учитывать [13, 14].

Для решения вышеперечисленных проблем предлагается новый метод обработки и анализа данных.

## 2. АНАЛИЗ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

Для анализа полученных данных был предложен специальный метод, учитывающий вероятности, погрешности и усреднения экспериментальных данных для получения релевантных результатов. Полученные данные анализировались на наличие изменений в количестве потребляемых воды и пищи, изменений массы тела, нарушений циркадных ритмов, изменений двигательной активности, изменения предпочтений в двигательной активности.

### 2.1. Выбор метода обработки

В связи с приведенными выше задачами для анализа данных был выбран  $U$ -критерий Манна–Уитни — непараметрический статистический критерий, используемый для сравнения двух независимых выборок по уровню какого-либо признака, измеренного количественно [15]. Метод основан на определении того, достаточно ли мала зона перекрещивающихся значений между двумя вариационными рядами, например между набором значений активности в экспериментальной и контрольной группах. Чем меньше значение критерия, тем вероятнее, что различия между значениями параметра в выборках достоверны. Значение  $U$ -критерия Манна–Уитни находится по формуле:

$$U = n_1 n_2 + \frac{n_x(n_x + 1)}{2} - T_x,$$

где  $n_1$ ,  $n_2$  — количество элементов в первой и второй выборке соответственно,  $T_x$  — большая из двух ранговых сумм, соответствующая выборке  $n_x$ . Полученное значение  $U$  сравнивается по таблице для избранного

уровня статистической значимости  $p$  с критическим значением  $U$  при заданной численности сопоставляемых выборок и делается вывод о наличии существенных различий [15]. Для нашей задачи уровень статистической значимости был выбран как  $p = 0.05$ .

$U$ -критерий Манна–Уитни был выбран как обладающий необходимыми качествами: он не требует нормального распределения сравниваемых совокупностей и одинаковых дисперсий, а также подходит для малых выборок. Недостатки данного критерия, такие как малое количество совпадающих данных и возможность сравнения только двух групп, не повлияли на анализ данных, так как в нашем исследовании требовалось сравнивать по две выборки, а отсутствие совпадающих данных обеспечивалось индивидуальностью живых объектов.

Данный метод является не только подходящим для решения нашей проблемы, но и рекомендуемым к использованию в подобных задачах [16]. Необходимо отметить, что  $U$ -критерий Манна–Уитни является широко используемым методом статистической обработки экспериментальных данных в медицине и биологии [17-20]. В связи с этим в данной работе был реализован предложенный метод на основе пакета SPSS Statistics, решающий задачу анализа и обработки данных эксперимента.

## 2.2. Первичный анализ данных

Перед началом обработки была выполнена подготовка данных — были построены и проанализированы графики веса, потребления и движения от времени. По результатам данного этапа одна из мышей была исключена из эксперимента, так как начала сильно терять в весе из-за болезни. Это повлияло на остальные показатели, такие как двигательная активность, и могло привести к неверным выводам в вопросе влияния наночастиц серебра. Также, были выполнены суммирование данных по различным временным промежуткам, соответствующим биологическим циклам животного, чтобы при сравнении выборок можно было говорить о сохранении различий при изменении временного шага.

## 2.3. Обработка $U$ -критерием Манна–Уитни

Анализ выполнялся при помощи распространенного для решения статистических задач пакета SPSS Statistics [21] на основе  $U$ -критерия Манна–Уитни для независимых выборок, усреднения данных по ансамблю и по времени (по 3, 4, 6 и 12 часов). Данные сравнивали попарно с учетом погрешностей (контрольную группу и группу, принимавшую наночастицы) в рамках одного периода по разным физиологическим показателям. Для каждого показателя был получен параметр  $p$ , который был сопоставлен с контроль-

ным значением, после чего был сделан вывод о значимости различий. Пример полученных значений параметра представлен в табл. 1 для группы «30 дней», сравнение проводилось по количеству потребляемой воды, пищи и оборотов колеса.

Далее, после обработки  $U$ -критерием Манна–Уитни был проведен конечный анализ, а именно было выполнено сравнение различий на предмет сохранения при изменении временного шага и сравнение разных возрастных групп.

По итогам анализа рассматриваемого эксперимента не было выявлено существенных различий у экспериментальных и контрольных групп животных в адаптации к новой обстановке, потреблению пищи и жидкости, двигательной активности и предпочтений в двигательной активности, а также суточных ритмах в системе для круглосуточного мониторинга PhenoMaster для групп 30, 60 и 180 дней. Пример графика представлен на рис. 1 — сравнение потребления жидкости и пищи, кручения колеса влево/вправо (по оси  $y$  отложены условные единицы).

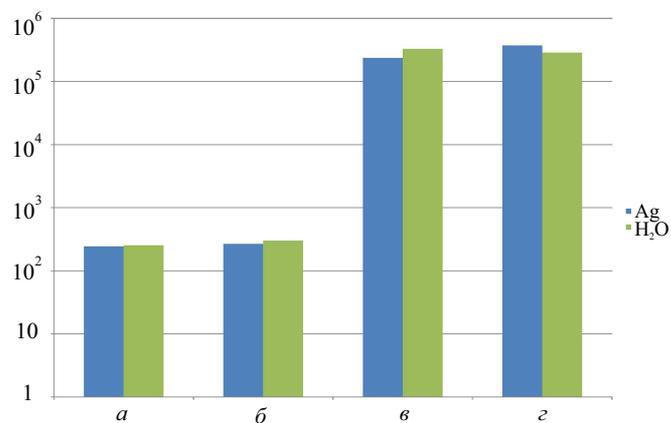


Рис. 1: Сравнение суммарных показателей потребления и двигательной активности (вращение колеса) для группы «60 дней»: *a* — суммарное потребление пищи, *b* — суммарное потребление воды, *v* — суммарное количество вращений колеса вправо, *z* — суммарное количество вращений колеса влево

Для группы 120 дней был выявлен надежный рост активности, так как наблюдалось сохранение существенных различий на различных временных интервалах (рис. 2-5). Погрешности не представлены на графиках и диаграммах, так как визуализация была выполнена исключительно в целях демонстрации значимых и незначимых различий, и при нанесении погрешностей на график различия становятся неочевидными. Значения  $p$  при анализе  $U$ -критерием Манна–Уитни были получены с учетом погрешностей, поэтому полученные результаты верны в пределах погрешностей.

Таблица I: Сравнения количества потребляемой жидкости, пищи и вращений колеса для контрольной и экспериментальной групп «30 дней»

Нулевая гипотеза	Значение	Решение
Распределение потребляемой жидкости одинаково для двух групп	0.645	Нулевая гипотеза принимается
Распределение потребляемой пищи одинаково для двух групп	0.442	Нулевая гипотеза принимается
Распределение количеств вращений колеса влево одинаково для двух групп	0.959	Нулевая гипотеза принимается
Распределение количеств вращений колеса вправо одинаково для двух групп	0.05	Нулевая гипотеза отклоняется
Распределение суммарных количеств вращений колеса одинаково для двух групп	0.328	Нулевая гипотеза принимается

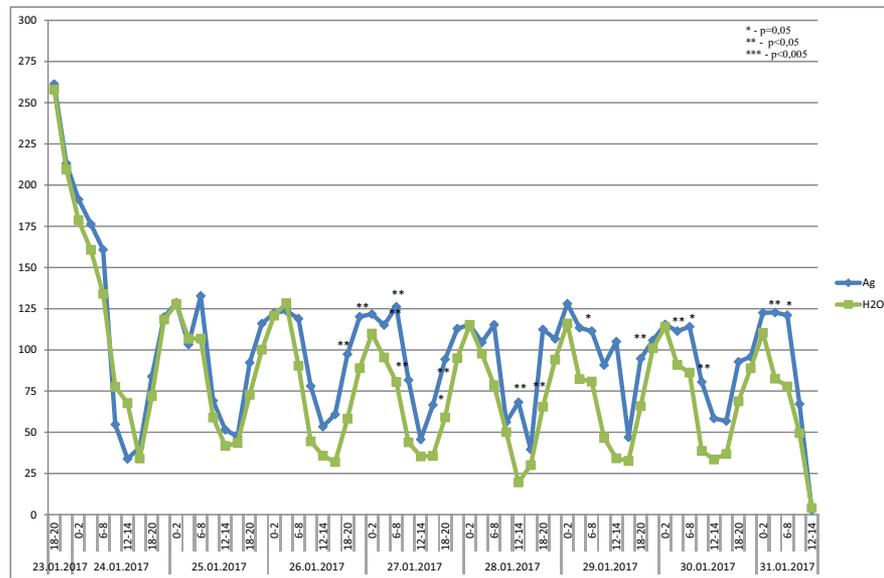


Рис. 2: Зависимость условных единиц активности от времени групп мышей «120 дней» в разбиении по 12 часов

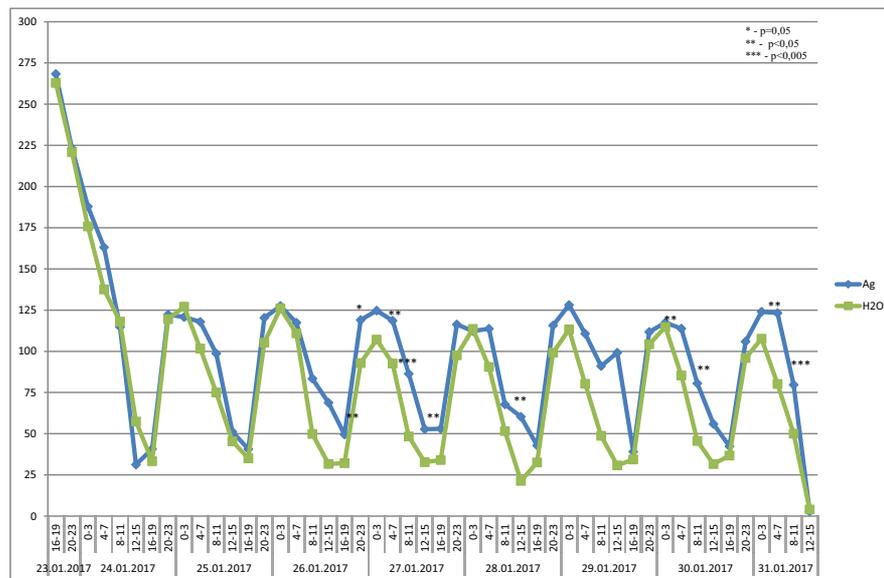


Рис. 3: Зависимость условных единиц активности от времени групп мышей «120 дней» в разбиении по 6 часов

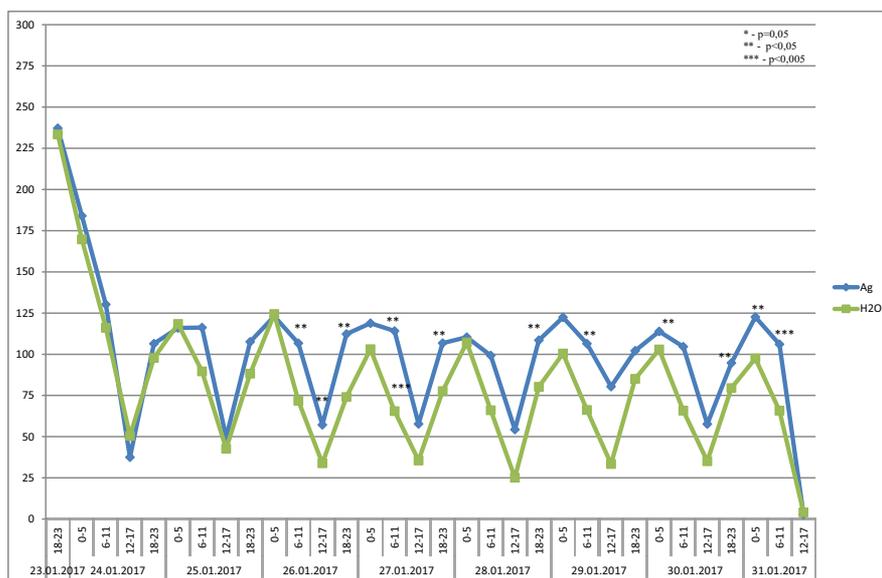


Рис. 4: Зависимость условных единиц активности от времени групп мышей «120 дней» в разбиении по 4 часа

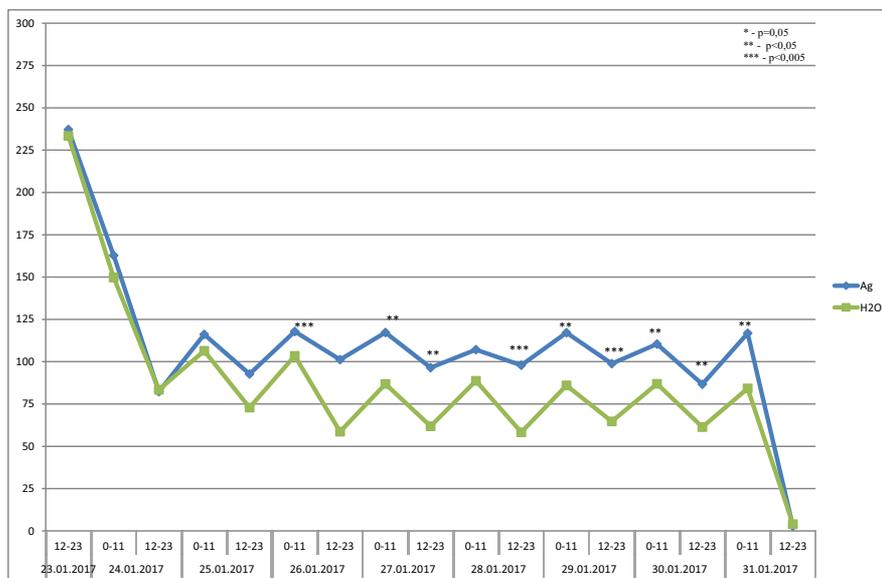


Рис. 5: Зависимость условных единиц активности от времени групп мышей «120 дней» в разбиении по 3 часа

**ЗАКЛЮЧЕНИЕ**

На данном этапе работы был предложен метод анализа большого количества экспериментальных данных, не требующий создания собственного программного обеспечения и легко подстраиваемый для решения подобных задач. Предложенный подход учитывает зависимость данных от времени и особенности живых организмов, такие как периодичность, индивидуальность, возраст, а также позволяет наглядно продемонстрировать различия между выборками.

По результатам анализа и интерпретации данных, были сделаны следующие выводы:

1. В связи с достоверностью полученных результатов можно сделать вывод о том, что предложенный метод применим к анализу подобного вида задач.
2. При интерпретации результатов нужно обращать внимание только на те различия, которые наследуются с периодичностью и сохраняются на нескольких временных разбиениях.

Автор надеется, что результаты данной работы могут сыграть значимую роль для решения вопросов анализа и интерпретации большого количества экспериментальных данных, полученных не только системой

PhenoMaster, но и в других системах мониторинга. Автор выражает благодарность сотрудникам РЦ нейрокогнитивных исследований НБИКС Курчатовского института: ведущему инженеру, канд. мед. наук

Долгову О.Н. за оказанную помощь и поддержку в работе, а также зоолаборанту–исследователю Бузову Н.Ю. за обеспечение идеального состояния экспериментальных животных и установки.

- [1] <https://www.tse-systems.com/products/behavior>
- [2] Bains R.S., Wells S., Sillito R.R., Armstrong J.D., Cater H.L., Banks G., Nolan P.M. J. Neurosci. Methods. 2018. **300**. P. 37.
- [3] Younis S., Schunke M., Massart J., Hjortebjerg R., Sundström E., Gustafson U., Björnholm M., Krook A., Frystyk J., Zierath J.R. et al. Proc. Natl. Acad. Sci. USA. 2018. **115**, N 9. E. 2048.
- [4] Barrington W.T., Wulfridge P., Wells A.E., Rojas C.M., Howe S.Y.F., Perry A., Hua K., Pellizzon M.A., Hansen K.D., Voy B.H. et al. Genetics. 2018. **208**, N 1. P. 399.
- [5] Drost H.-G., Paszkowski J. Bioinformatics. 2017. **33**, N 8. P. 1216.
- [6] Watson M., Thomson M., Risse J., Talbot R., Santoyo-Lopez J., Gharbi K., Blaxter M. Bioinformatics. 2015. **31**, N 1. P. 114.
- [7] Бухтеева И.А., Долгов О.Н., Анциферова А.А., Копеева М.Ю. XXV Международная конференция студентов, аспирантов и молодых ученых по фундаментальным наукам Международный молодежный форум «Ломоносов-2018», секция «физика», подсекция «медицинская физика», сборник тезисов докладов. 2018. С. 268.
- [8] Antsiferova A.A., Kopeeva M.Yu., Kashkarov P.K., Kovalchuk M.V. Proceedings Book 3rd International Symposium on Nanoparticles-Nanomaterials and Applications. 2018. P. 178.
- [9] Ильин В.П. Acta Biomedica Scientifica. 2011. N 5. С. 157.
- [10] Fay M.P., Proschan M.A. Stat Surv. 2010. **4**. P. 1.
- [11] Siepka S.M., Takahashi J.S. Methods Enzymol. 2005. **393**. P. 230.
- [12] Hut R.A., Pilorz V., Boerema A.S., Strijkstra A.M., Daan S. PLoS One. 2011. **6**, N 3. E. 17527.
- [13] Shoji H., Takao K., Hattori S., Miyakawa T. Mol Brain. 2016. **9**, N 11. P. 1.
- [14] Starr M.E., Saito H. J Gerontol A Biol Sci Med Sci. 2012. **67**, N 10. P. 1043.
- [15] Rivas-Ruiz R., Moreno-Palacios J., Talavera J.O. Rev Med Inst Mex Seguro Soc. 2013. **51**, N 4. P. 414.
- [16] Altman D.G., Gore S.M., Gardner M.J., Pocock S.J. BMJ. 1983. **286**. P. 1489.
- [17] Schwerk A., Jaskaia R. Peer J. 2018. **6**. E. 4657.
- [18] Han B., Zhu F., Yu H., Liu S., Zhou J. Sci Rep. 2018. **8**. E. 5956.
- [19] Graziano M., Benito R., Planas J.V., Palstra A.P. BMC Dev. Biol. 2018. **18**. E. 10.
- [20] Гублер Е.В., Генкин А.А. Применение непараметрических критериев статистики в медико-биологических исследованиях. Медицина, 1973.
- [21] Dembe A.E., Partridge J.S., Geist L.C. BMC Health Serv Res. 2011. **11**, N 252. P. 1.

## Data analysis of the automatic round-the-clock monitoring system PhenoMaster based on the Mann-Whitney U test

I. A. Bukhteeva

Department of General Physics and Molecular Electronics, Faculty of Physics, Lomonosov Moscow State University, Moscow 119991, Russia  
E-mail: iabukhteeva@mail.ru

PhenoMaster is a unique modular autonomic platform. It provides scientists from different research areas modern technology for fully automated and synchronized metabolic, behavioral and physiological monitoring: indirect gas calorimetry, food and fluid intake, locomotor activity, learning, memory, etc. Data recording can be performed for any given time interval, for example, every second and for any given period. The output is a huge amount of data, which is very difficult to analyze and interpret using a personal computer. This work is devoted to the solution of this problem.

Data analysis was carried out on the basis of a comprehensive study the effect of oral injection of biologically active additives based on silver nanoparticles on the behavioral, cognitive and physiological functions of mammals. Physiological features of animals were recorded in the system of automatic monitoring PhenoMaster and for analyzing the obtained data, a unique method was developed. It takes into account measurement errors, time averaging and ensemble averaging, as well as features of living organisms to obtain relevant results. The obtained data were analyzed for changes in the amount of consumed water and food, changes in body weight, disorders of circadian rhythms, changes in motor activity and its preferences. The analysis was based on the Mann-Whitney u test for independent samples.

PACS: 87.50.sg, 07.05.Kf, 29.85.+c

Keywords: PhenoMaster, analysis of experimental data, data processing, statistics, Mann–Whitney U test, nonparametric test.

Received 03 May 2018.

**Сведения об авторе** Бухтеева Ирина Алексеевна — студентка; e-mail: iabukhteeva@mail.ru.