

Нохрин Д. Ю. Малоизвестные ранговые критерии как альтернатива параметрическим методам статистической обработки биомедицинских данных / Образование магистров: проблемы и перспективы развития. Тезисы докладов VIII Всероссийской конференции. Челябинск, Изд-во Челяб. гос. ун-та, 2025. С. 107-115.

УДК 519.257, 378.14

**МАЛОИЗВЕСТНЫЕ РАНГОВЫЕ КРИТЕРИИ КАК АЛЬТЕРНАТИВА
ПАРАМЕТРИЧЕСКИМ МЕТОДАМ СТАТИСТИЧЕСКОЙ ОБРАБОТКИ
БИОМЕДИЦИНСКИХ ДАННЫХ**

Нохрин Д. Ю.

Челябинский государственный университет, Челябинск, Россия

nokhrin8@mail.ru

В настоящее время в биологических и медицинских научных исследованиях применяются как параметрические, так и непараметрические статистические методы анализа данных. Корректное использование первых предполагает нормальное распределение признаков в генеральной совокупности, из которой извлекается выборка, что требует либо специальной статистической проверки, либо применения нормализующих преобразований исходных данных с последующим контролем эффективности их работы. Наиболее распространёнными преобразованиями для количественных признаков являются угловые преобразования для частот (например, фи-преобразование) и степенные преобразования для положительно асимметрично распределённых признаков – площадей, скоростей, концентраций в широком диапазоне, численностей организмов (например, преобразование квадратного корня, логарифмическое преобразование или адаптивное к данным преобразование Бокса – Кокса). Поскольку биомедицинские данные, как правило, не подчиняются нормальному распределению, а использование преобразований требует от исследователя большей квалификации, широкое распространение получили непараметрические ранговые аналоги классических параметрических тестов. Так, например, сравнение мер положения двух выборок проводят не критерием Стьюдента, а критерием Манна – Уитни, мер положения нескольких выборок – не однофакторным дисперсионным анализом, а критерием Краскела – Уоллиса, поиск связей двух признаков – не линейной корреляцией Пирсона, а ранговыми корреляциями Спирмена или Кендалла и. т.п. Большинство из таких альтернатив классическим параметрическим методам были разработаны в середине XX века, тогда как более современные техники остаются малоизвестными. Цель данной работы – привлечь внимание к двум ранговым техникам анализа данных, которые должны быть востребованными в научных исследованиях биомедицинского профиля, в том числе – студентами, магистрантами и аспирантами в ходе подготовки квалификационных работ.

1. Критерий Бруннера – Мунцеля (Brunner–Munzel test), как альтернатива критериям Стьюдента и Манна – Уитни. Критерий Уилкоксона – Манна – Уитни заслуженно получил очень широкое распространение для сравнения двух независимых выборок по количественным или качественным порядковым признакам. Данный критерий свободен от распределения и имеет очень высокую мощность: его питмановская асимптотическая эффективность равна $3/\pi$ или 95,5%. Однако даже в отношении столь известного метода распространены заблуждения. Первое – что он сравнивает медианы выборок; указания на это имеются даже в литературе по прикладной статистике и руководствах к статистическим пакетам. В действительности, в отличие от медианного критерия Муда (Mood's median test), он сравнивает не медианы, а функции двух распределений [6], что обычно не принципиально с точки зрения интерпретации результатов сравнения мер положения, но может иметь особенности в некоторых частных случаях. Второе заблуждение – его безусловность, тогда как в действительности он требует равенства дисперсий в популяциях, откуда извлекаются выборки. То-есть, строго говоря, критерий Манна – Уитни является ранговым аналогом классического критерия Стьюдента для сравнения средних значений независимых выборок с равными дисперсиями. Поскольку исторически ранговые методы получили популярность именно как более простые и вычислительно менее трудоёмкие, проверка этого требования не вошла в практику статистического анализа данных, а в некоторых грамотных статистических руководствах требование равенства дисперсий упростили до требования одинакового типа распределения в выборках [2], которое можно оценить визуально (по гистограммам) не прибегая к расчётам. Вместе с тем, в области параметрической статистики был накоплен большой массив информации, указывающей на то, что дисперсии в выборках, например, больных и здоровых лиц обычно различаются, а потому в практике биомедицинских исследований чаще стал применяться вариант *t*-критерия Стьюдента в модификации Уэлча (Welch's *t*-tes), том числе после преобразования исходных данных. В этом подходе реализована идея своеобразного штрафа за несоблюдение требований равенства дисперсий; на практике это выражалось в получении дробных степеней свободы (degree of freedom, *df*), ручной расчёт которых с последующей интерполяцией *p*-значений по таблицам распределения Стьюдента для целочисленных *df* были крайне трудоёмкими. Однако с распространением персональных компьютеров данный метод прочно вошёл в исследовательскую практику, и в настоящее время подход Уэлча по умолчанию используется в ряде статистических техник, где требуется сравнение средних значений: например, при ранжировании генов для построения вулканного графика (Volcano plot) в биоинформатике или выбор точки отсечения количественного признака для выработки правил «если-то» (if-then rules) в технике снижения многофакторной размерности (multifactor dimensionality reduction, MDR) при моделировании взаимодействия генов.

Критерий Бруннера – Мунцеля был разработан в 2000 г. именно как ранговый аналог *t*-критерия Стьюдента в модификации Уэлча двумя сотрудниками Института медицинской статистики Университета в Гёттингене (Германия): Эдгаром Бруннером (Edgar Brunner) и Ульрихом Мунцелем (Ullrich Munzel) [5]. Транслитерация названия критерия в русском языке не устоялась, имеются варианты написания «критерий Бруннера – Мюнцеля», а на англоязычных ресурсах он произносится как «Бруннер Манзел тест». Учитывая, немецкое происхождение критерия правильное произносить и

транслитерировать данный эпоним как немецкий. Критерий проверяет нулевую гипотезу о том, что для случайно выбранных значений X и Y из двух генеральных совокупностей вероятность того, что X будет больше Y , равна вероятности того, что Y будет больше X . Данный критерий не требует равенства дисперсий и рекомендуется к применению в случаях, когда традиционно использовался критерий Манна – Уитни без проверки требования равенства дисперсий рангов или одинаковой формы распределения в выборках. Название недавней статьи Дж. Карча звучит более категорично: «Психологи **должны** использовать тест Бруннера – Мунцеля вместо U -теста Манна – Уитни в качестве непараметрической процедуры по умолчанию» [9]. Учитывая собственный опыт и опыт других исследователей логично распространить данное требование также на биологов и медиков.

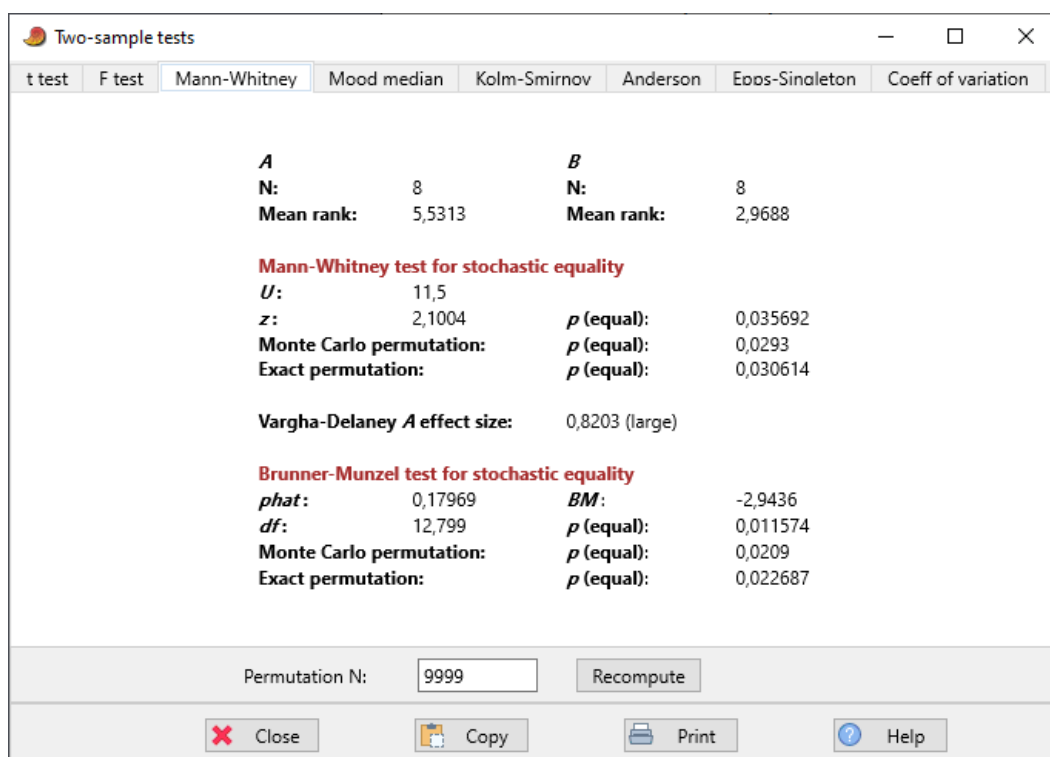


Рис. 1 – Результат вычисления критерия Бруннера – Мунцеля в пакете PAST

На сегодняшний день критерий Бруннера – Мунцеля реализован в двух пакетах для кроссплатформенной программно-статистической среды R: *lawstat* (<https://cran.r-project.org/web/packages/lawstat/index.html>) и *brunnermunzel* (<https://cran.r-project.org/web/packages/brunnermunzel/index.html>), причём во втором случае имеется также перестановочный (permutation) вариант расчёта p -значения. В статье [10] авторы этого подхода приводят результаты обширного моделирования, которые показывают, что для очень малых и очень разных размеров выборок, что обычно для медицинских и биологических исследований, перестановочный вариант хорошо работает для данных из нескольких основных распределений. Также критерий Бруннера – Мунцеля включен в бесплатные статистические пакеты, разработанные И. П. Гайдышевым: статистический пакет *StatAnt*, начиная с версии 4.0.0 для Linux и Windows (<https://sourceforge.net/projects/statant/>) и статистический пакет *Criterial* (настройка для LibreOffice Calc или Apache OpenOffice Calc), начиная с версии 2.1.0 для Linux, UNIX и

Windows (<https://sourceforge.net/projects/criterial/>). Студентам биологического и экологического факультета ЧелГУ также просто идти в ногу со временем в анализе данных, поскольку в базовом статистическом пакете PAST [8], используемом на практических занятиях по дисциплинам «Основы биометрического анализа и планирование эксперимента» и «Методы математической статистики в экологии и природопользовании» программ бакалавриата и «Компьютерные технологии в биологии. Математическое моделирование биологических процессов», критерий Бруннера – Мунцеля был недавно реализован (версии 4.16. и старше). Причём, как и во многих других модулях этого пакета, имеется возможность расчёта p -значения перестановочной техникой Монте-Карло (для любых объёмов выборок) и точным рандомизационным методом (для выборок до $n=25$ включительно). Для расчета нужно выделить данные и проследовать по пути: Univariate – Two sample tests – Two sample tests (F, t...) – Закладка Mann-Whitney. На рис. 1 приведён скриншот работы модуля пакета для данных примера из нашего Лабораторного практикума ([4], стр. 96-97).

2. Порядковая регрессия (Ordinal Regression) с моделями кумулятивных связей (Cumulative link models, CLM), как альтернатива методам дисперсионного анализа. Начиная с работ Рональда Фишера – автора дисперсионного анализа (Analysis of variance, ANOVA) и «гения, который практически в одиночку создал основы современной статистической науки» [7], различные модели ANOVA стали неотъемлемым инструментом анализа данных во всех научных областях. Разработанная в 1970-х годах теория общих линейных моделей (General Linear Models, GLM) обобщила и частично вытеснила широкий спектр традиционных статистических методов поиска различий и зависимостей (все варианты t -критерия, дисперсионного, регрессионного и ковариационного анализов и др.). К настоящему времени в прикладных статистических пакетах за привычными названиями таких методов в меню могут стоять расчётные алгоритмы GLM, которые позволяют корректно обрабатывать непропорциональные дисперсионные комплексы, наборы с частично пропущенными блоками данных, рассчитывать смешанные модели для нестандартных экспериментальных планов и т.п.). С появлением в те же годы обобщённых линейных моделей (Generalized Linear Models, аббревиатура также GLM и понимается из контекста) к количественным показателям добавились качественные номинальные, что расширило область применения моделей также на логлинейный анализ, логистическую регрессию и регрессию Пуассона. Слабым звеном всех этих методов была работа с порядковыми данными, активное развитие методов для обработки которых началось в 1980-е годы после разработки Маккалахом модели пропорциональных шансов (proportional odds model of McCullagh) – частного случая большого семейства моделей, появившихся позже и получивших название моделей кумулятивных связей (Cumulative link models, CLM). Разработка гибких CLM-моделей и их доступная программная реализация позволяют в настоящее время получить полноценные ранговые аналоги GLM-моделей. Рассмотрим ниже использование CLM-модели в качестве функционального рангового аналога двухфакторного дисперсионного анализа.

До недавнего времени ранговый дисперсионный анализ при наличии двух факторов можно было провести в 2 этапа с использованием критерия Фридмана (Friedman test), который является ранговым аналогом однофакторного дисперсионного анализа с

повторными наблюдениями, а также двухфакторного дисперсионного анализа для рандомизированных блоков. При этом на первом этапе оценивались различия между выборками по первому фактору, на втором – по второму, а взаимодействие факторов оставалось невыраженным и оба раза входило в ошибку анализа. Такой аналог модели ANOVA для главных эффектов является неполноценным, поскольку в случае реального наличия взаимодействия факторов именно оно представляет главный теоретический и практический интерес и может приводить к неверной интерпретации вследствие парадокса Симпсона. Поэтому в случае ненормального распределения признака на практике обычно использовали подходящее преобразование данных, которое устраняло эффект шкалы и делало распределение ошибки модели приближенным к нормальному, а далее использовали параметрический дисперсионный анализ. Именно так, например, мы поступили в исследовании, где сравнивали концентрации цитокинов в плазме крови выборки здоровых детей и детей со средней и тяжёлой формами внебольничной пневмонии с учётом возраста [3]. В табл. 1 воспроизведены результаты дисперсионного анализа из этой статьи (табл. 2, стр. 948) по интерлейкину IL-28A, выполненного на предварительно преобразованных по Боксу – Коксу данных (Box–Cox transformation), а в верхней части рис. 2 – соответствующий график взаимодействия факторов, выполненный в логарифмическом масштабе после наивной ретрансформации (naïve retransformation) средних значений и границ 95%-ных доверительных интервалов из преобразованной шкалы в исходную (рис. 1, стр. 949).

Таблица 1 – Результаты двухфакторного дисперсионного анализа различий в концентрации IL28A в сыворотке крови у детей с пневмониями и группе сравнения

Источник изменчивости	Сумма квадратов <i>SS</i>	Степени Свободы <i>df</i>	Средний квадрат <i>MS</i>	<i>F</i> -критерий	<i>p</i>
Главные эффекты					
Заболевание	771,070	2	385,535	29,41	<0,001
Возраст	66,278	3	22,093	1,69	0,173
Взаимодействие факторов					
Заболевание × Возраст	111,287	6	18,548	1,41	0,213
Ошибка	1743,592	133	13,110	–	–

Покажем далее, как можно провести такой анализ с использованием не степенного, а рангового преобразования исходных данных, т.е. с использованием модели кумулятивной связи. Для анализа нами использовалась среда R с установленным пакетом ordinal (<https://cran.r-project.org/web/packages/ordinal/index.html>). Исходные данные были представлены таблицей MS Excel для 145 детей с колонками Group (метки 1, 2, 3), Age (1, 2, 3, 4) и последующими колонками с концентрациями цитокинов, включая колонку IL28A. Ниже приведён программный код R для проведения анализа с небольшими комментариями.

```
library(ordinal) # подключение пакета ordinal
Копирование таблицы с данными в буфер обмена.
mydata<-read.table("clipboard", dec=".", h=T) # создание набора данных mydata из
содержимого буфера обмена с указанием на запятую в качестве десятичного разделителя и
на наличие заголовка в первой строке (header=TRUE)
mydata$Group<-as.factor(mydata$Group) # перекодирование цифровых
выборок в метки фактора
mydata$Age<-as.factor(mydata$Age) # перекодирование цифровых обозначений
категорий возраста в метки фактора
mydata$Rank.f<-factor(rank(mydata$IL28A), ordered = TRUE) # создание колонки с
переменной Rank.f с заменой концентраций IL28A рангами и перекодировкой этих рангов
в метки фактора (таково требование модели)
model <- clm(mydata$Rank.f ~ Group + Age + Group:Age, data = mydata,
threshold="symmetric") # подгонка CLM-модели с рангами из колонки Rank.f в качестве
зависимой переменной (отклика) и меток для выборок и возрастов – в качестве
независимых факторов (регрессоров).
anova(model, type="III") # аналог дисперсионного анализа полученной model
После выполнения последней команды появляются, следующий результат:
Type III Analysis of Deviance Table with Wald chi-square tests
```

	Df	Chisq	Pr(>Chisq)
Group	2	39.169	3.123e-09 ***
Age	3	19.852	0.0001822 ***
Group:Age	6	15.571	0.0162534 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Видно, что для оценки главных факторов и их взаимодействия использовалась статистика хи-квадрат Вальда и все оцененные эффекты были статистически значимыми. Полученный результат сведён в табл. 2.

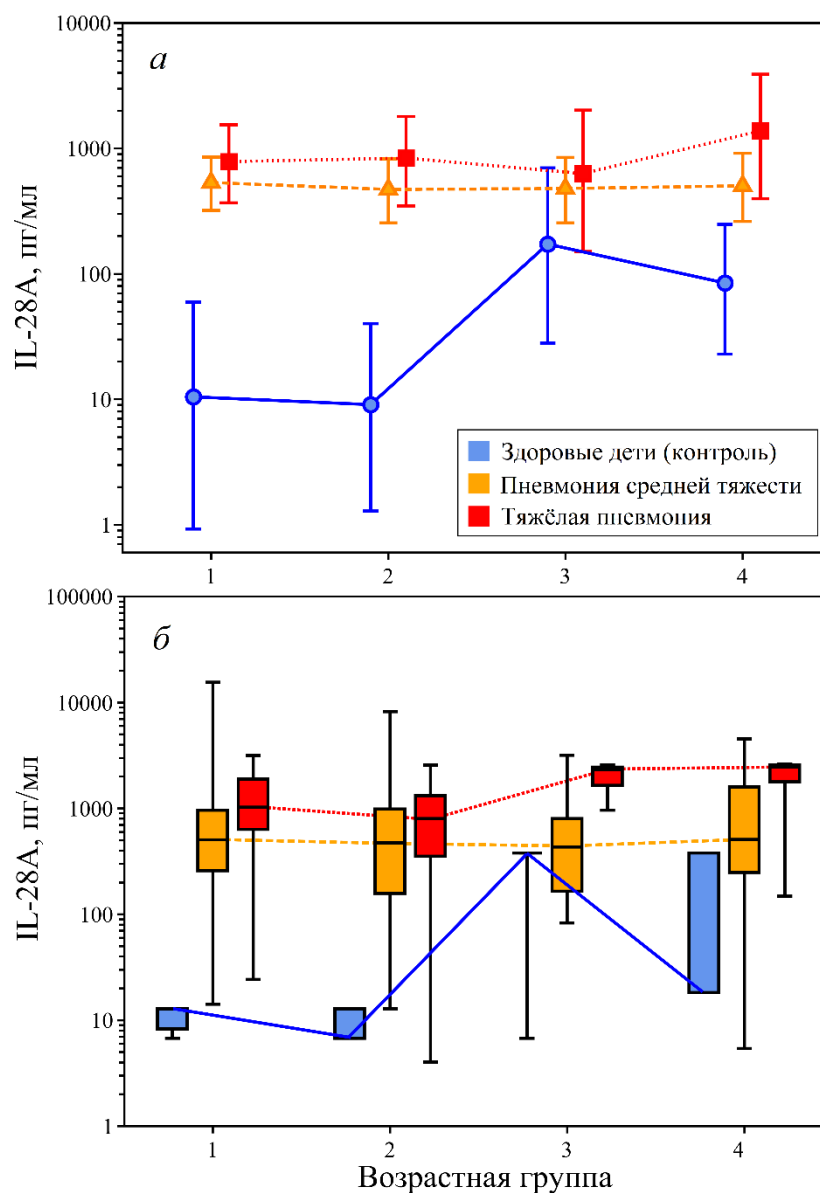


Рис. 2 – Концентрация IL-28A в сыворотке крови у детей с пневмониями и в группе сравнения в зависимости от возраста: а) средние значения с 95% ДИ по результатам двухфакторного дисперсионного анализа, б) коробчатый график. Пояснения – в тексте.

Таблица 2 – Результаты модели кумулятивной связи анализа различий в концентрации IL28A в сыворотке крови у детей с пневмониями и группе сравнения

Источник изменчивости	Степени свободы <i>df</i>	Статистика хи-квадрат χ^2	<i>p</i>
Заболевание	2	39,17	<0,001
Возраст	3	19,85	<0,001
Заболевание × Возраст	6	15,57	0,016

Таким образом, использование рангового аналога двухфакторного дисперсионного анализа позволило статистически доказать не только различия между группами детей, но и между возрастными, а также подтвердить ($\chi^2_{(6)} = 15,57$; $p=0,016$) – а значит и получить возможность обсуждать – существование взаимодействия факторов. Оно заключалось в том, что у детей с пневмониями возрастные изменения не были выражены, в то время как в группе здоровых детей имелась динамика: в возрасте 1-3 лет (возрастная группа 1) и 4-7 лет (возрастная группа 2) уровни IL-28A были близки, далее к 8-12 годам увеличивались и к 13-18 годам несколько снижались. Как видно из рис. 2-б, если соединить медианы коробчатой диаграммы (Box-and-Whisker plot) линиями профилей, то эти профили оказываются почти идентичными таковым для рис. 2-а (малое число наблюдений в группе сравнения ($n=28$) привело к совпадению медиан и квартилей, что выразилось в нестандартном виде коробов и усов).

Важно отметить, что как показал выборочный анализ данных по другим цитокинам из этого исследования, далеко не всегда близкие к статистически значимым эффекты стали таковыми в модели порядковой регрессии, наблюдалась и обратная ситуация. В связи с этим возникает вопрос: а какую технику анализа тогда предпочтительнее использовать? У автора нет однозначного ответа на этот вопрос, поскольку выбор из нескольких близких статистических техник часто диктуется личными предпочтениями, контекстом исследования, традициями в данной области науки или в данном конкретном журнале. Поэтому единственное, что мы стараемся делать – избегать мозаичного и придерживаться единообразного подхода к анализу данных. То-есть, если мы собираемся приводить в таблицах медианы с квартилями и строить коробчатые диаграммы, то логично продолжать использовать порядковую статистику и для решения других задач в исследовании, включая порядковую корреляцию, регрессию и ранговые техники редукции данных с обобщением.

Заключение. Использование рассмотренных ранговых методов – критерия Бруннера – Мунцеля и порядковой регрессии с моделями кумулятивных связей – позволят магистрантам и аспирантам анализировать данные своих исследований более современными и корректными статистическими техниками и повысить качество своих квалификационных работ. Отметим также, что всё большее число современных методов анализа и визуализации научных данных требует освоения программно-статистической среды R. Применительно к теме данной статьи интересующимся можно рекомендовать ознакомиться со статьёй [1] и другими публикациями этого авторского коллектива, а также активнее осваивать англоязычные онлайн-ресурсы, среди которых многие носят именно обучающий характер.

Список литературы:

1. Егошин, В. Л. Порядковая регрессия в программной среде R – краткие рекомендации для магистрантов и докторантов по специальности «Медицина» и «Общественное здравоохранение» / В. Л. Егошин, Н. В. Саввина, А. М. Гржибовский // *West Kazakhstan Medical Journal*. – 2019. – Т. 61, № 3. – С. 145-153.
2. Закс, Л. Статистическое оценивание / Л. Закс ; пер. с нем. В. Н. Варыгина ; под ред. Ю. П. Адлера, В. Г. Горского. М. : Статистика, 1976. 598 с.
3. Изюрова, Н.В. Цитокиновый профиль при внебольничной пневмонии у детей / Н. В. Изюрова, А. Ю. Савочкина, А. Н. Узунова, Д. Ю. Нохрин // *Медицинская иммунология*. – 2022. – Т. 24, № 5. – С. 943–954. DOI: 10.15789/1563-0625-CPI-2538

4. Нохрин, Д. Ю. Лабораторный практикум по биостатистике / Д. Ю. Нохрин. Челябинск : Изд-во Челяб. гос. ун-та, 2018. 289 с.
5. Brunner, E. The nonparametric Behrens–Fisher problem: Asymptotic theory and a small-sample approximation / E. Brunner, U. Munzel. // Biometrical Journal. – 2000. – V. 42, no 1. – P. 17–25. DOI: 10.1002/(SICI)1521-4036(200001)42:1<17::AID-BIMJ17>3.0.CO;2-U
6. Encyclopedia of statistical sciences / S. Kotz, C.B. Read, N. Balakrishnan, B. Vidakovic, N.L. Johnson. – 16 Volume Set, 2nd Edition. – Hoboken : John Wiley & Sons, Inc., 2005. – 9686 p.
7. Hald, A. A history of mathematical statistics / A. Hald. – New York : Wiley. 1998. – 738 p.
8. Hammer, Ø. PAST: Paleontological statistics software package for education and data analysis / Ø. Hammer, D. A. T. Harper, P. D. Ryan // Palaeontologia Electronica. – 2001. – № 1. – P. 1–9.
9. Karch, J. D. Psychologists should use Brunner–Munzel's instead of Mann–Whitney's U test as the default nonparametric procedure / J. D. Karch // Advances in Methods and Practices in Psychological Science. – 2021. – V. 4, no 2. – P. 1–14. DOI: 10.1177/2515245921999602.hdl:1887/3209569
10. Neubert, K. A studentized permutation test for the non-parametric Behrens–Fisher problem / K. Neubert, E. Brunner // Computational Statistics & Data Analysis. – 2007. – V. 51, no 10. – P. 5192–5204. DOI: 10.1016/j.csda.2006.05.024

УДК 577.3

ГЕНОТОКСИЧЕСКИЕ ЭФФЕКТЫ У МЫШЕЙ ПРИ КОМБИНИРОВАННОМ ДЕЙСТВИИ ГАММА-ИЗЛУЧЕНИЯ И ИМПУЛЬСНОГО МАГНИТНОГО ПОЛЯ

Орехова К.С.

Челябинский государственный университет, Челябинск, Россия

ksyusha.orehova@list.ru

Введение

Современные исследования в области биологии и медицины все больше обращаются к вопросам воздействия различных физических факторов на живые организмы. Одной из актуальных тем является изучение генотоксических эффектов. В последние десятилетия научное сообщество проявляет растущий интерес к комбинированному воздействию этих факторов, что обусловлено их широкой встречаемостью в быту, промышленности, медицине [4, 5].

Цель данной работы заключается в изучении генотоксических эффектов у мышей при комбинированном воздействии гамма-излучения и импульсного магнитного поля. В рамках исследования были сформулированы следующие задачи:

1. Рассмотреть влияние гамма-излучения на генетический материал мышей;
2. Изучить эффекты, вызываемые импульсным магнитным полем;
3. Оценить частоту появления микроядер в периферической крови у самцов мышей при облучении в различных дозах гамма-облучения + импульсным магнитным полем.

Также, в настоящее время остро стоит проблема необходимости проведения эффективного, быстрого и недорогого генетического мониторинга с целью оценки