

УДК 613+311.21

<https://doi.org/10.23946/2500-0764-2023-8-1-119-131>

ТЕХНОЛОГИИ СТАТИСТИЧЕСКОГО АНАЛИЗА МЕДИЦИНСКИХ ДАННЫХ: ПЕРВИЧНЫЙ АНАЛИЗ ДАННЫХ, СРАВНЕНИЕ ГРУПП

ГУДИНОВА Ж. В., ДЕМАКОВА Л. В.*

ФГБОУ ВО «Омский государственный медицинский университет» Министерства здравоохранения Российской Федерации, г. Омск, Россия

Резюме

В публикации представлены материалы по проведению наиболее распространённых, несложных и информативных процедур первичного статистического анализа медицинских данных. Подчеркивается значение непараметрических показателей описательной статистики (медиана, процентиля) и критериев сравнения групп, важность и технология расчётов референсных интервалов, процедур сортировки и ранжирования данных, необходимость использования гистограммы как самостоятельного аналитического инструмента. Выделены этапы анализа данных. Материал иллюстрирован графиками и скриншотами, облегчающими

восприятие текста и самостоятельное освоение статистических процедур в формате ППП Statistica. На рисунках приведены пошаговая инструкция первичного анализа показателей физического развития детей, гистограммы показателей инвалидности детей в регионах России.

Ключевые слова: описательная статистика, анализ данных, гистограмма, процентиль.

Конфликт интересов

Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

Источник финансирования

Собственные средства.

Для цитирования:

Гудина Ж.В., Демакова Л.В. Технологии статистического анализа медицинских данных: первичный анализ данных, сравнение групп. *Фундаментальная и клиническая медицина*. 2023;8(1): 119-131. <https://doi.org/10.23946/2500-0764-2023-8-1-119-131>

*Корреспонденцию адресовать:

Демакова Лидия Владимировна, 644099, Россия, г. Омск, ул. Ленина, д. 12, E-mail: dlv1210@yandex.ru dlv1210@yandex.ru
© Гудина Ж.В. и др.

LECTURE

STATISTICAL ANALYSIS OF MEDICAL DATA: DESCRIPTIVE STATISTICS AND GROUP COMPARISON

ZHANNA V. GUDINOVA, LIDIA V. DEMAKOVA *

Omsk State Medical University, Omsk, Russian Federation

For citation:

Zhanna V. Gudina, Lidia V. Demakova. Statistical analysis of medical data: descriptive statistics and group comparison. *Fundamental and Clinical Medicine*. (In Russ.). 2023;8(1): 119-131. <https://doi.org/10.23946/2500-0764-2023-8-1-119-131>

*Corresponding author:

Dr. Lidia V. Demakova, 12, Lenina Street, Omsk, 644099, Russian Federation, E-mail: dlv1210@yandex.ru
© Zhanna V. Gudina, et al.

English ►

Abstract

Here we describe the most common and informative procedures for the primary statistical analysis of medical data. We emphasize the importance of using histogram, normality tests, median, percentiles, confidence intervals, ranking procedures, and as analytical tools. The material is illustrated with step-by-step graphical explanation of applying commercially avail-

able software for fulfilling these tasks. The lecture is designed for biomedical students and clinicians.

Keywords: descriptive statistics, data analysis, histogram, percentile.

Conflict of Interest

None declared.

Funding

There was no funding for this project.

Несмотря на большой перечень показателей первичного анализа данных, их высокую эффективность в отношении получения новой полезной информации об изучаемом объекте [1-19], используется обычно весьма ограниченный перечень показателей первичного анализа данных или основных статистик – среднее арифметическое, стандартное отклонение, ошибка среднего. Отечественные исследователи обычно игнорируют медиану, процентиля, гистограмму как самостоятельный аналитический инструмент, асимметрию и эксцесс как характеристики распределения. То есть в отсутствие применения всего перечня простых, но информативных инструментов первичного анализа данных, теряется много важных выводов, не говоря уже о решении вопроса о выборе параметрических или непараметрических методов анализа в зависимости от характера распределения данных.

Начиная разбор первичного анализа данных, следует усвоить понятие *типа данных* или показателей: 1) количественные, получаемые при измерениях (например, данные о росте, массе тела, температуре, времени и т. п.), изменяющиеся от 0 до бесконечности, 2) порядковые данные, соответствующие их местам в возрастающем или убывающем порядке (например, это степени болезни – 1, 2, 3, 4 и т.д.), 3) качественные данные, представляющие собой определенные свойства элементов выборки или популяции, которые нельзя измерить (мальчики – девочки, болен – здоров, курит – не курит и т.п.). Следует заранее, перед началом сбора базы данных, решить вопрос о том, каким образом те или иные данные будут собираться. При этом нужно понимать, что порядковые и тем более качественные данные гораздо менее информативны. Всегда следует стремиться к получению количественных данных. Например: молодой или пожилой – это данные качественные либо порядковые (может быть еще категория «зрелый возраст», например). Но лучше всего

в базе данных фиксировать возраст человека в годах (или дату рождения и дату исследования), так как потом можно разделить группу на нужные категории, а вот наоборот сделать нельзя. Кроме того, для порядковых и качественных данных перечень методов статистического анализа ограничен, по сравнению с количественными данными, и значит, получение новых полезных результатов ограничено.

Первичный анализ данных. Статистический анализ количественных данных в медицинских исследованиях начинается, как правило, с оценки распределения и с описательной статистики.

Описательная статистика – комплекс базовых приемов анализа данных вариационного ряда, дающий наиболее общее представление о распределении той или иной характеристики в изучаемой совокупности.

Вариационный ряд – это упорядоченное изображение реально существующего распределения наблюдений в группе по величине признака.

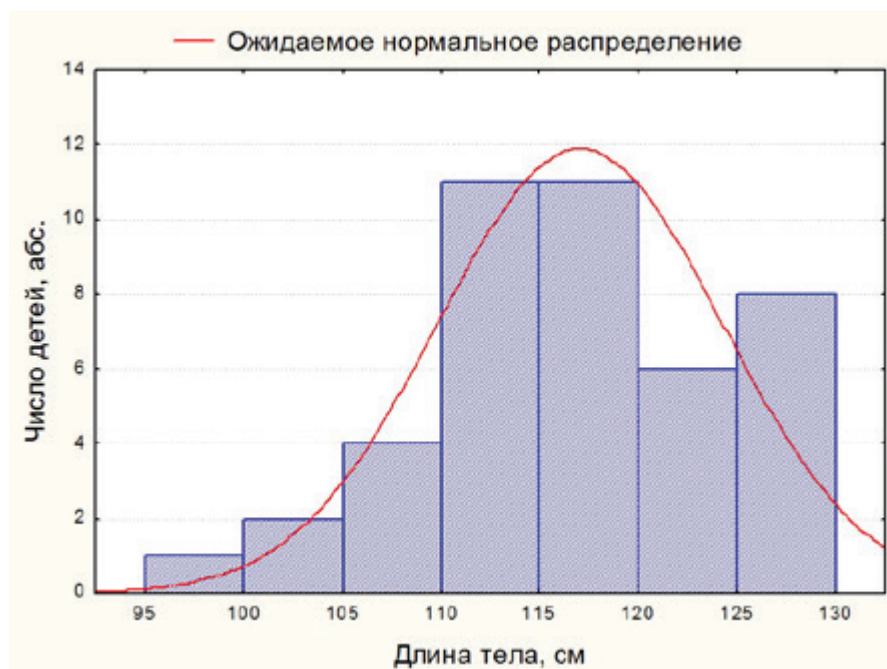
Гистограмма – это график частотного распределения, проще говоря, это столбиковая диаграмма, характеризующая сколько элементов изучаемой совокупности (или наблюдений – детей, животных, регионов), по оси ОУ вошло в тот или иной диапазон значений изучаемого показателя по оси ОХ (**рисунок 1**).

Основные параметры описательной статистики (распределения): среднее арифметическое, ошибка среднего, медиана (срединное значение вариационного ряда), мода (наиболее частое значение ряда), стандартное отклонение (мера вариабельности данных в ряду), минимум (минимальное значение в ряду), максимум (максимальное значение в ряду), процентиля, счет (количество членов ряда), доверительный интервал, коэффициент вариации др.

Этапы анализа распределения: 1) расчёт параметров описательной статистики (распределения); 2) графический анализ графика

распределения (гистограммы) с визуальной оценкой смещения центра распределения ряда вправо или влево (асимметрия распределения), вверх-вниз (эксцесс распределения) по отношению к красной линии (рисунок 1 А); 3) рас-

чет критериев нормальности распределения (χ^2 , критерий Колмогорова-Смирнова, W критерия Шапиро-Уилка, значение $p < 0,05$ полученного критерия указывает на значимое отклонение распределения от нормального) (рисунок 1 Б).



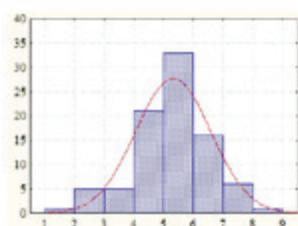
А.

Рисунок 1.

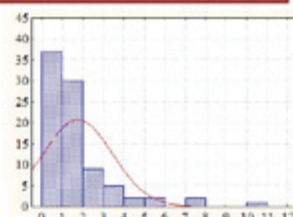
Информативная ценность гистограммы (выполнено в пакете прикладных программ Statistica): А – распределение детей-дошкольников по показателю Длина тела, Б – распределение регионов России по показателям инвалидности детей вследствие разных болезней, на 10 тыс. детей, с указанием уровня значимости p полученного критерия Шапиро-Уилка.

Figure 1. Information value of the histogram (made in the Statistica software): А – distribution of preschoolers by body length, В – distribution of disability indicators across different regions of Russian Federation (per 10,000 children), Shapiro-Wilk test.

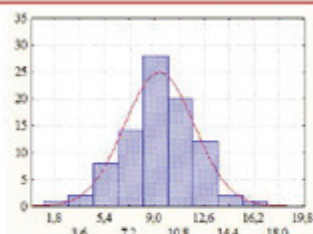
Нормальное распределение ($p > 0,05$)



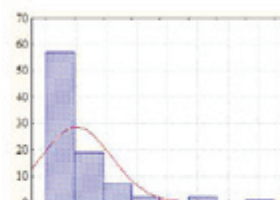
Распределение отличается от нормального ($p < 0,05$)



Новообразования ($p = 0,327$)



Инфекционные и паразитарные болезни ($p = 0,00025$)



Болезни эндокринной системы ($p = 0,50090$)

Отдельные состояния, ...в перинатальном периоде ($p = 0,00070$)

Б.

Гистограмма. Для оценки распределения существует ряд процедур, в том числе, визуальный анализ гистограммы. Если гистограмма принимает колоколообразную форму («колокол Гаусса»), то распределение данного признака отвечает закону о нормальном распределении, которое в ППП Statistica на графике гистограммы автоматически изображается линией красного цвета (**рисунок 1**). При этом крайние значения показывают редко встречающиеся значения, а при приближении к середине кривой частота встречаемости значений в наблюдаемой группе увеличивается. Это следует понимать так, что людей среднего роста в популяции всегда больше, чем высоких и низких. Однако с другими показателями (например, патологическими сдвигами) так бывает далеко не всегда. Например, распределение массы тела во взрослой популяции всегда будет асимметрично, с более длинной правой ветвью, поскольку крайнее максимальное значение массы тела человека в популяции будет намного больше отличаться от центра распределения по сравнению с крайним минимальным.

Для примера на **рисунок 1А** приведена гистограмма, характеризующее распределение показателей длины тела в группе детей дошкольного возраста. По рисунку можно сделать вывод о том, что распределение признака визуально отличается от нормального (то есть от линии ожидаемого нормального распределения). То есть в дальнейшем следует использовать непараметрические показатели и критерии. Кроме того, гистограмма позволяет получить еще некоторый объем информации:

рост детей в группе колеблется в диапазоне от 95 до 130 см;

наибольшее количество детей (по 11 человек) вошли в два диапазона: от 110 до 115 см и от 115 до 120 см;

довольно много детей имеют максимальные показатели роста (для данной группы): 8 человек вошли в диапазон от 125 до 130 см.

И так далее. Подобная информация может представлять большой интерес, как теоретический, так и практический. Так, можно легко решить вопрос, сколько мебели того или иного номера по ГОСТу¹ следует приобрести для данной группы детей:

¹ *цит. по СанПиН 1.2.3685-21. Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания.*

мебель группы 0 (для детей ростом 850-1000 мм) – 1 комплект,

мебель группы 1 (для детей ростом 1000-1150 мм) – 17 комплектов,

мебель группы 2 (для детей ростом 1150-1200 мм) – 25 комплектов.

На **рисунок 1 Б** распределение регионов России по показателям инвалидности детей вследствие разных болезней. Судя по форме графиков и приведенным показателям уровня значимости p критерия Шапиро-Уилка, нормальное распределение имеют показатели инвалидности детей в регионах России по классам новообразования и болезни эндокринной системы: графики симметричные, с одной медианой в центре распределения, соответствуют линии ожидаемого нормального распределения, $p > 0,05$. Поэтому можно сделать вывод о минимальном воздействии внешних влияний при формировании инвалидности детей по этим болезням. Внешними влияниями здесь могут быть факторы как среды обитания в регионах, так и организационные, такие как показания к определению инвалидности, более или менее четкие. И, напротив, по классам инфекционных и паразитарных болезней и отдельных состояний, возникающих в перинатальном периоде очевидно отличие распределения от нормального: правосторонняя асимметрия, с длинной правой ветвью и с отсутствием левой ветви, высокий эксцесс (высота столбиков диаграммы над линией нормального распределения), $p < 0,05$. В таком случае следует предположить влияние внешних факторов на формирование инвалидности детей вследствие указанных болезней. Следует установить регионы с высокими показателями и задуматься о причинах подобных состояний. Так, на гистограмме Инфекционные болезни² видно, что порядка 70 российских регионов вошли в диапазоны минимальных значений – от 0 до 2 на 10 тыс. Но в одном регионе показатели порядка 10–11 на 10 тыс. детского населения, то есть в 5–10 раз выше, чем в основной массе российских территорий. Подобная ситуация и по классу Отдельных состояний в перинатальном периоде: на большинстве территорий страны (около 80 регионов) показатели очень низкие – 0–2 на 10 тыс., а в одном – 7 на 10 тыс. Явно нужно поставить вопрос о причинах подобных различий.

² *большой частью это туберкулез.*

Следует подчеркнуть, что подобные нюансы распределения практически невозможно (долго, трудно) выявить без графика, анализируя столбики цифр.

Процентили и медиана. Очень важно с точки зрения первичного анализа данных, а особенно для решения задачи разработки референсных интервалов (нормативов), понятие процентилей. *Процентиль (P)* – это одна из числовых характеристик распределения, процентиль делит совокупность на 100 равных частей так, что P0 – это минимальное значение, а P100 – максимальное. Первый процентиль P1 – это такая точка или значение в этой совокупности (считая от меньшего значения вверх), ниже которой находится 1% всех случаев. 50-й процентиль P50 (или медиана) – такая точка или значение, ниже которой находятся 50% всех признаков и т. д. *Дециль* делит совокупность на десять частей (например, третий дециль P30 – это точка, ниже которой находятся 30% случаев), *квантиль* P20 – на пять частей, *квартиль* P25 – на четыре части. Любой из них может быть использован для определения дисперсии вокруг медианы и для определения границ референсных интервалов, хотя децильные (P10 и P90) и квартильные (P25 и P75) ранги наиболее часто встречаются в литературе.

Общеизвестно, что среднее арифметическое, медиана и мода в случае нормального распределения совпадают, а при распределении, отличном от нормального, существенно различаются. Однако, к сожалению, большинство отечественных исследователей для описания центра распределения и вариабельности изучаемого признака используют среднее арифметическое и стандартное отклонение, не углубляясь в изучение характера распределения, что приводит к ошибочным результатам статистического описания данных. Более правильно для описания центра распределения использовать медиану, а для характеристики вариабельности признака – процентили (P10, P25, P75, P90), тем более что медиана более приближается к интуитивному пониманию среднего, чем часто критикуемое среднее арифметическое (особенно часто критикуется уровень средней заработной платы).

Доверительный интервал – параметрический показатель, характеризующий диапазон значений, в котором выборочное среднее \bar{M} является серединой этого диапазона. Если предположить, что $\alpha = 0,05$, то нужно определить ту часть стандартной нормальной кривой,

которая равна $(1 - \alpha)$, или 95 процентам. Это значение равно $\pm 1,96$. Следовательно, доверительный интервал, определяется по формуле: $ДИ = \bar{M} \pm 1,96 (\delta / \sqrt{n})$, где \bar{M} – среднее значение совокупности, δ – стандартное отклонение, n – объем выборки (совокупности). В MS Excel доверительный интервал определяется как $(\bar{M} \pm \text{ДОВЕРИТ})$.

Коэффициент вариации V – параметрический показатель, позволяющий судить о разбросе данных и степени однородности изучаемой совокупности (когорты, группы). Рассчитывается по формуле: $V = \delta / \bar{M} * 100\%$, где V – коэффициент вариации, δ – стандартное отклонение, \bar{M} – среднее значение.

Если коэффициент вариации не превышает 10%, то можно говорить о слабом разбросе, а совокупность считать однородной; пределах 10-30% – разброс средний, совокупность – средняя по однородности. При $V > 30\%$ разброс вариант считают большим, а совокупность неоднородной. В этом случае следует искать причины подобной неоднородности.

На **рисунке 2** приведена пошаговая инструкция первичного анализа базы данных по исследованию физического развития детей.

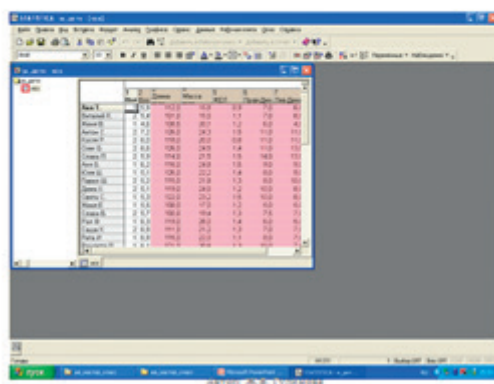
Выводы для дальнейшего анализа по **рисунку 2**: распределение данного показателя отличается от нормального, значит, параметрические критерии использовать нельзя. То есть, для описания данных следует использовать медиану (а не среднее арифметическое) и процентили (а не стандартное отклонение и ошибки среднего). А для сравнения двух независимых групп следует использовать непараметрический критерий Манна-Уитни (а не критерий Т-Стьюдента).

Таким образом, описательная статистика обладает самостоятельным значением, так как позволяет не только получить первое, общее представление об изучаемой совокупности, а также решить вопрос выбора параметрических или непараметрических методов дальнейшего анализа данных, но и, в случае исследования какого-то нового показателя (например, артериального давления новорожденного с низкой массой тела), позволяет разработать его референсные интервалы (или границы статистической нормы) в целях последующего использования полученных нормативов на практике. Это достойная задача медицинского научного исследования, но, к сожалению, она редко ставится отечественными исследователями,

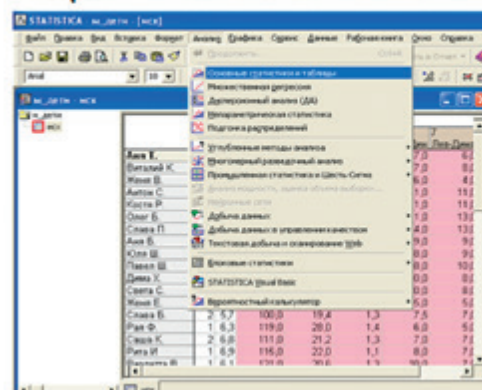
Рисунок 2.
Пошаговая инструкция первичного анализа данных в ППП Statistica (описательная статистика) на примере изучения физического развития дошкольников (шаги 1-8).

Figure 2.
Step-by-step instructions for analysing sample distribution and descriptive statistics in Statistica software. Example: studying the physical development of pre-schoolers (steps 1-8).

Шаг 1: открыть файл в программе Statistica



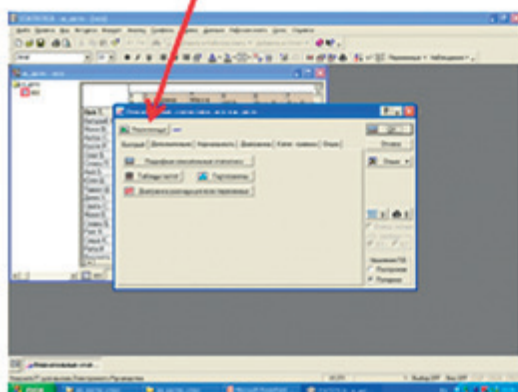
Шаг 2: выбрать в меню Анализ опцию Основная статистика /



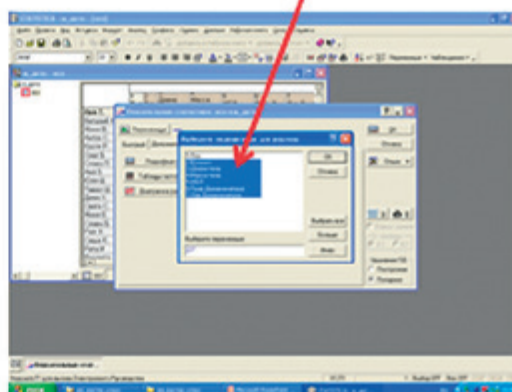
Автор: Ж.В. Гуркина

31

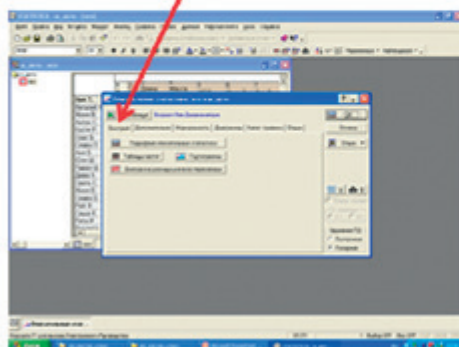
ШАГ 3: выбрать окно Переменные



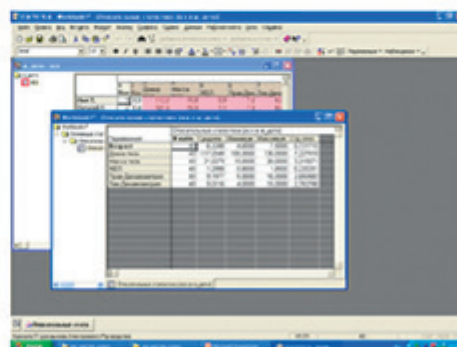
ШАГ 4: в окне Переменные выбрать переменные №№ 2-7



ШАГ 5: выбрать Быстрый (анализ).



ШАГ 6: Машина выдаст вам среднее значение, максимум, минимум, стандартное отклонение по показателям Возраста, Длины тела, массы тела, ЖЕЛ и динамометрии. Проанализировать полученный результат.



ШАГ 6: фрагмент – одним нажатием кнопки в ППП Statistica можно получить огромное количество результатов (43 – это количество детей). В MS Excel такое невозможно: их можно получить только по отдельности. Вот оно – решение наших задач:

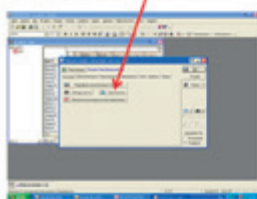
1. Определить средние значения показателей физического развития (рост, масса тела и др.)
2. Определить минимальные и максимальные показатели среди детей.

Описательные статистики (исх в м_дети)

Переменная	Описательные статистики (исх в м_дети)	Н набл.	Среднее	Минимум	Максимум	Стд. откл.
Возраст		43	6,2283	4,6000	7,5000	0,721715
Длина тела		43	117,0349	100,0000	130,0000	7,227918
Масса тела		43	21,0279	15,0000	30,0000	3,315571
ЖЕЛ		43	1,2983	0,8000	1,8500	0,220251
Прав-Динамометрия		43	9,1977	5,0000	16,0000	2,650560
Лев-Динамометрия		43	9,0115	4,0000	15,0000	2,782788

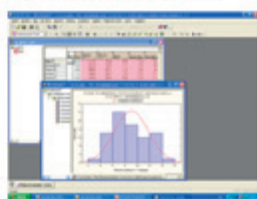
Затем следует оценить характер распределения данных, чтобы решить вопрос об использовании в дальнейшем параметрики или непараметрики

ШАГ 7: выбрать
Нормальность – Критерий
Шapiro-Уилка – Гистограммы



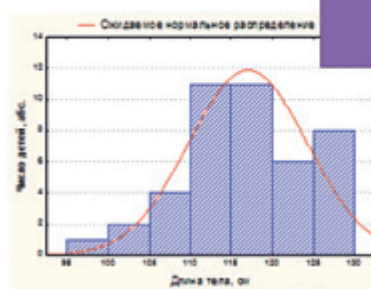
Автор: Ж.В. Гуринкин

ШАГ 8: проанализировать
полученные результаты и
графики (гистограммы)



Автор: Ж.В. Гуринкин

ШАГ 8: оценка характера распределения данных длины тела



Вывод: распределение
отличается от
нормального,
асимметрично

Автор: Ж.В. Гуринкин

возможно, из-за недостаточного знакомства с возможностями и процедурами описательной статистики.

Расчет референсных интервалов. В ходе описания результатов исследования, для того чтобы врач мог применить ваши результаты на практике, следует приводить не $M \pm m$ (M – среднее арифметическое, m – ошибка среднего), а $M \pm S$ (S – стандартное отклонение). В случае приведения ошибки m такое применение почти невозможно: ведь каждый конкретный человек вовсе не обязан иметь значения, максимально приближенные к средней арифметической. В случае же приведения стандартного отклонения S читатель становится понятно, что в этот интервал входят 68% наблюдений ($P84-P16$).

Расчёты референсных интервалов (или интервалов статистической нормы) проводятся на основании вариационного ряда с распределением показателей от минимального значения до максимального и с учётом от характера распределения (таблица 1).

Как видно из таблицы 1, в случае нормального распределения для расчёта границ нормы признака следует применить метод сигмальных отклонений (и тогда центром распределения будет среднее значение (M_{cp}), а границы ранговых коридоров будут рассчитываться в стандартных отклонениях (δ). В случае распределения, отличного от нормального, следует применить метод процентилей (и тогда центром распределения будет медиана (Me или $P50$), а границы ранговых коридоров будут рассчитываться в значениях процентилей (P). На практике в качестве границ референсных интервалов часто применяют $P25-P75$, этот интервал уже, в него входят всего 50% наблюдений, а не 68% как в случае $P16-P84$ (таблица 1).

Более распространены в настоящее время референсные интервалы, рассчитанные по методу процентилей, и описываются они таким образом: $Me (P25-P75)$ или $Me (P25; P75)$, где Me – медиана. Когда авторы игнорируют характер распределения признака и приводят по традиции интервалы параметрические, $M + S$, то нижняя граница такого

Таблица 1.
Границы статистической нормы признака, рассчитанные разными методами.

Table 1.
Reference range of the trait as calculated by different methods.

Номер рангового коридора Number	Границы / Range		Название рангового коридора / Category
	В случае нормального распределения / Normal distribution	В случае распределения, отличного от нормального / Non-normal distribution	
1	$< (M - 2\delta)$	до $P_{2,5}$	Очень низкие значения / Very low values
2	$-1\delta - 2\delta$	$P_{2,5} - P_{16}$	Низкие значения / Low values
3	$M \pm \delta$	$P_{16} - P_{84}$	Средние значения / Average values
4	$+1\delta + 2\delta$	$P_{84} - P_{97,5}$	Высокие значения / High values
5	$> (M + 2\delta)$	выше $P_{97,5}$	Очень высокие значения / Very high values

интервала становится отрицательным значением, ниже 0, чего не может быть по природе явления. Или наоборот, верхняя граница референсного интервала превышает максимально возможное значение признака. Всё это грубые ошибки, не позволяющие использовать полученные результаты.

Процедуры сортировки и ранжирования данных также можно отнести к первичному анализу данных. Они дают представление о том, какой элемент совокупности находится на первом и последних местах в изучаемой совокупности, позволяют построить наглядный график убывания, например, избыточной массы тела у детей в изучаемой совокупности ев-

ропейских стран (**рисунок 3**) и т.п. Эти процедуры осуществляются в MS Excel.

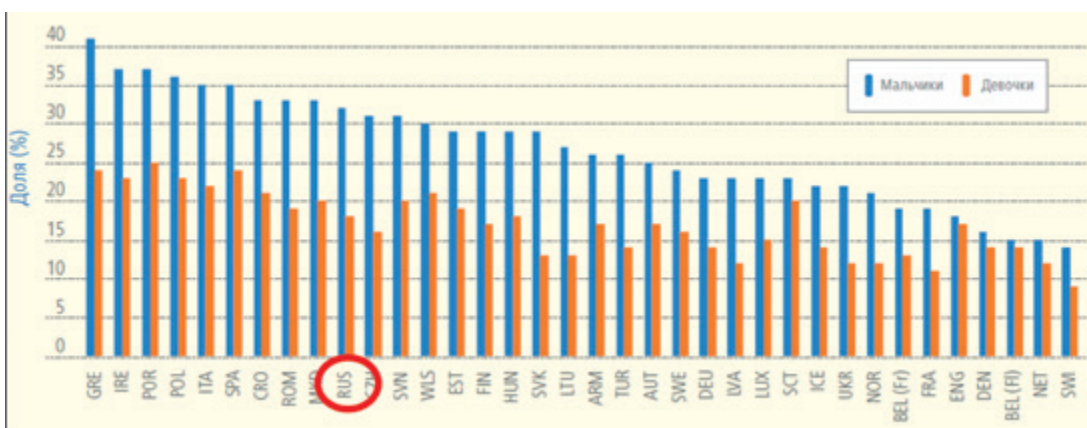
По графикам на рисунках 3 можно сделать следующие выводы:

- 1) если по показателям избыточной массы тела и ожирения среди взрослого населения Россия находилась в середине списка европейских стран, то среди детей Россия входила в первую десятку стран, это тревожная тенденция;
- 2) во всех странах среди мальчиков уровни избыточной массы тела и ожирения выше, чем среди девочек, и разница нередко была существенной;
- 3) визуально различия по изучаемому признаку здоровья населения (максимум – минимум)

Рисунок 3.
Распространённость избыточной массы тела и ожирения среди населения европейских стран, %: А – взрослое население, лица старше 18 лет, 2014 г.; Б – дети в возрасте 11 лет, 2009/2010 гг. [5].



Figure 3.
Prevalence of overweight and obesity among the population of European countries, %: A – adult population > 18 years of age, 2014; B – 11-years-old children, 2009/2010 [5].



между странами более сильно выражены среди детей – почти втрое, тогда как среди взрослых меж-
страновые различия выражены гораздо менее.

Методы сравнения групп являются классическими и наиболее распространенными в медицинских исследованиях. Сравнение данных с указанием статистической значимости различий – основа статистического анализа данных. Таким образом можно оценивать, насколько случайны различия между группами, следует ли принимать во внимание различия, являются ли они следствием влияния каких-то важных факторов и нужно ли принимать какие-то меры.

Существует масса способов, выбор которых основывается на цели и задачах исследования. Способы сравнения планируются обычно в начале исследования в ходе разработки дизайна исследования. Очень часто в медицинских исследованиях стоит задача сравнения групп (например, основной и контрольной, либо одной и той же группы – до и после лечения). Групп может быть две (two), три и более (multiple). Различают малую (менее 30), среднюю (от 30 до 100) и большую выборку (более 100 испытуемых).

Для выбора критерия сравнения групп важно различать характер выборки. Существуют выборки *независимые*, то есть не связанные между собой (мужчины и женщины, больные и здоровые), и *зависимые*, например, полученные в двух опытах, на одной и той же группе испытуемых (до и после воздействия).

Все критерии основаны на применении т.н. нулевой гипотезы (о том, что различия между группами не значимы, не существенны) и определении её справедливости. Если вероятность получить наблюдаемые или более сильные различия в повторных выборках мала, то нулевая гипотеза отвергается, то есть различия между группами значимы. В медико-биологических исследованиях уровень значимости принимается $p < 0,05$. В особых случаях уровень значимости различий может быть принят как $p < 0,01$. Некоторые авторы выделяют ещё *статистически значимую тенденцию* (при $p = 0,05-0,1$), что позволяет не отвергать важный результат. Распространённая ошибка: в одной публикации в одном случае приводить $p < 0,05$, а в другом – $p < 0,01$ или $p < 0,001$. Уровень значимости принимается единый для всей работы или оговаривается заранее, в каких случаях увеличивается точность исследования. По современным требованиям, при описании полученных результатов в тексте научной работы следует указывать не просто $p < 0,05$, а точное его значе-

ние, например, $p = 0,045$, как выдаёт его аналитическая программа. Формат « $p < 0,05$ » используется в отечественной медицинской науке просто по инерции, с тех давних пор, когда приходилось вручную выбирать уровень значимости в табличных приложениях к руководствам по статистическому анализу данных [10].

Один из основных методов оценки различий между группами – *дисперсионный анализ* основанный на сравнении разброса выборочных средних в группах с разбросом значений внутри группы с помощью дисперсии и расчета критерия F. Разновидностью дисперсионного анализа в случае сравнения двух групп является *t-критерий*, который также имеет варианты – сравнение независимых и зависимых групп. Следует помнить, что широкое распространение t-критерия в отечественных медицинских исследованиях неправомерно, применение этого критерия требует нормального распределения данных и соблюдения ряда других условий, которые в совокупности случаются весьма редко. Тем более, что современные пакеты прикладных программ включают его непараметрические аналоги, использование которых и более корректно, и просто.

Наиболее частые задачи в медицинских исследованиях могут быть решены с помощью следующих непараметрических критериев, основанных на оценке статистической значимости различий между группами:

критерий Краскела-Уоллиса (H) – применяется для сравнения между несколькими независимыми группами, тремя и более (как альтернатива параметрическому дисперсионному анализу, ANOVA);

критерий Манна-Уитни (U) – применяется для сравнения распределений в двух независимых группах наблюдений (например, при сравнении роста мальчиков и девочек), аналог Т-критерия Стьюдента;

критерий Вилкоксона (W) – применяется для сравнения двух зависимых групп (до и после воздействия фактора, периода времени, учебного года и проч.).

Все критерии непараметрические, то есть могут применяться без предварительной оценки характера распределения, в чем и состоит их достоинство. Недостаток – они более грубые по сравнению с параметрическими аналогами.

Расчёт этих критериев можно проводить на базе STATISTICA с применением опций Непараметрические методы (критерии Манна-Уитни, Краскела-Уоллиса, Уилкоксона). Все опции предпо-

лагают выбор группирующей переменной, а также графический анализ данных (рисунки 4–5).

В качестве примера применения критерия Манна-Уитни в гигиенических исследованиях продолжим статистическую обработку базы данных по исследованию физического развития детей. Ранее нами была поставлена задача сравнить длину тела у мальчиков и девочек. Расчёты показывают, что среднее значение роста у мальчиков составляет 117 см, а у девочек – 116 см. На первый взгляд, мальчики выше. Проверим это утверждение на основании определения уровня статистической значимости различий в продолжении пошаговой инструкции (рисунок 2, 4). В начале анализа на базе ППП Statistica следует выбрать группирующую переменную. В нашем случае это качественный

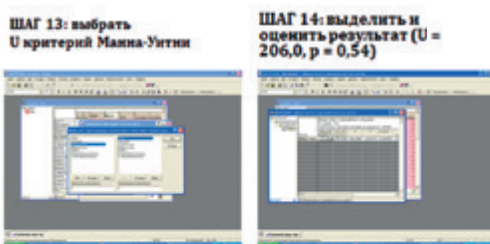
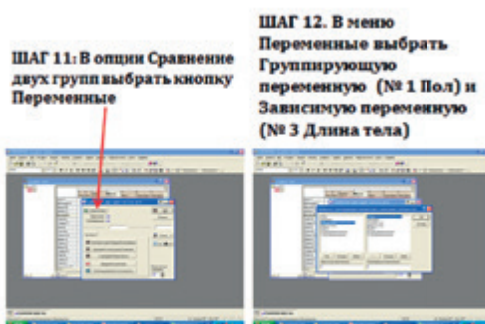
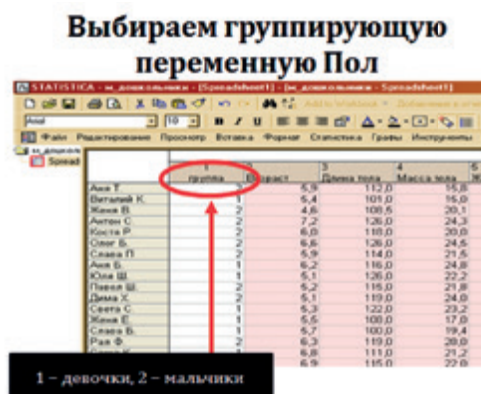
признак, разделение группы детей по полу: 1 – мальчики, 2 – девочки.

Проанализируем графики, полученные в ходе шагов 15, 16 (рисунок 4), – диаграмму размаха (ее еще называют *коробочной диаграммой* или *бокс-диаграммой*) и категоризованную гистограмму. На рисунке 5 приведены окончательные результаты сравнения групп, которые следует приводить в публикации, – или один из графиков, или оба, так как они дают разную информацию. Однако в любом случае по сравнению со столбиковой диаграммой предпочтителен коробочный график – он имеет массу преимуществ при сравнении групп: даёт больше информации, позволяет получить больше выводов и нагляднее показывает отсутствие (наличие) различий между группами.

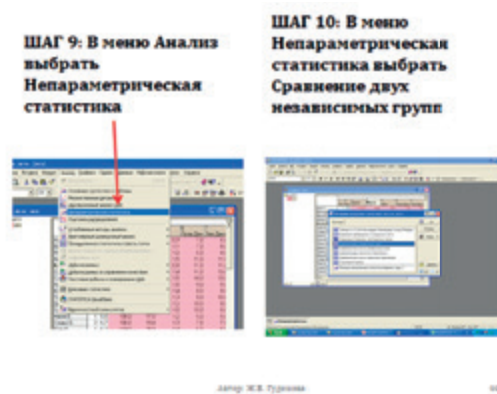
Рисунок 4.

Пошаговая инструкция сравнения групп с помощью критерия Манна-Уитни в ППП Statistica на примере изучения физического развития дошкольников (шаги 9–16).

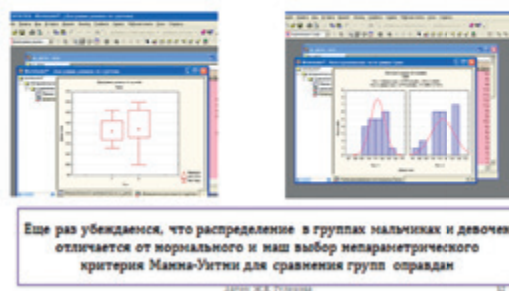
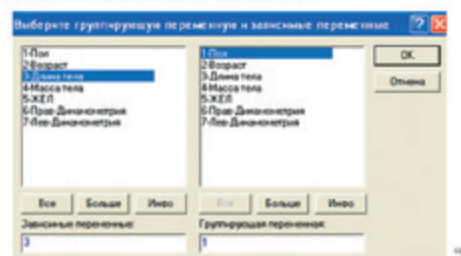
Figure 4. Step-by-step instructions for comparing groups using Mann-Whitney U-test in Statistica software. Example: studying the physical development of preschoolers (steps 9–16).

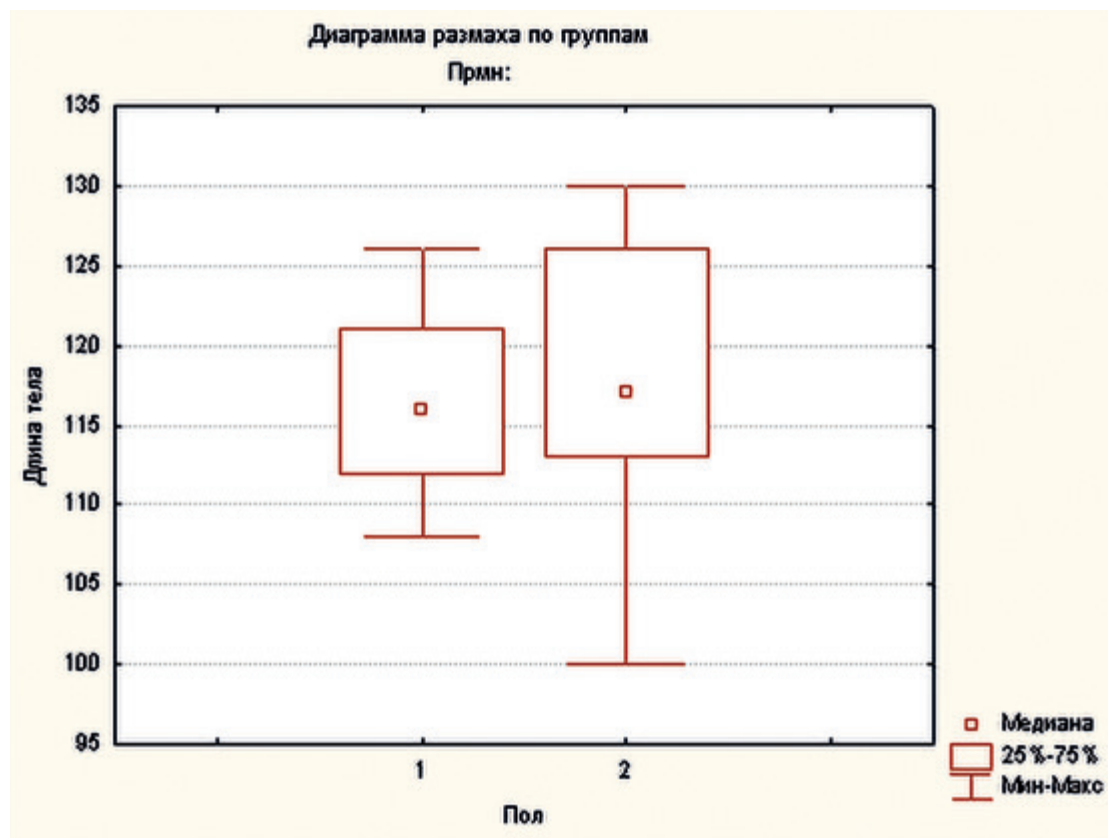


Ответьте на вопрос: значимы ли различия длины тела у мальчиков и девочек? Если заданный уровень $p \leq 0,05$, а полученный $p = 0,54$?



Для этого метода принципиальна группирующая переменная. И их может быть много: группировать свою выборку можно по-разному – и по полу, и по возрасту, и в какой детский сад ходит, и в каком районе живет, и больной, и здоровый... И так далее, и тому подобное.

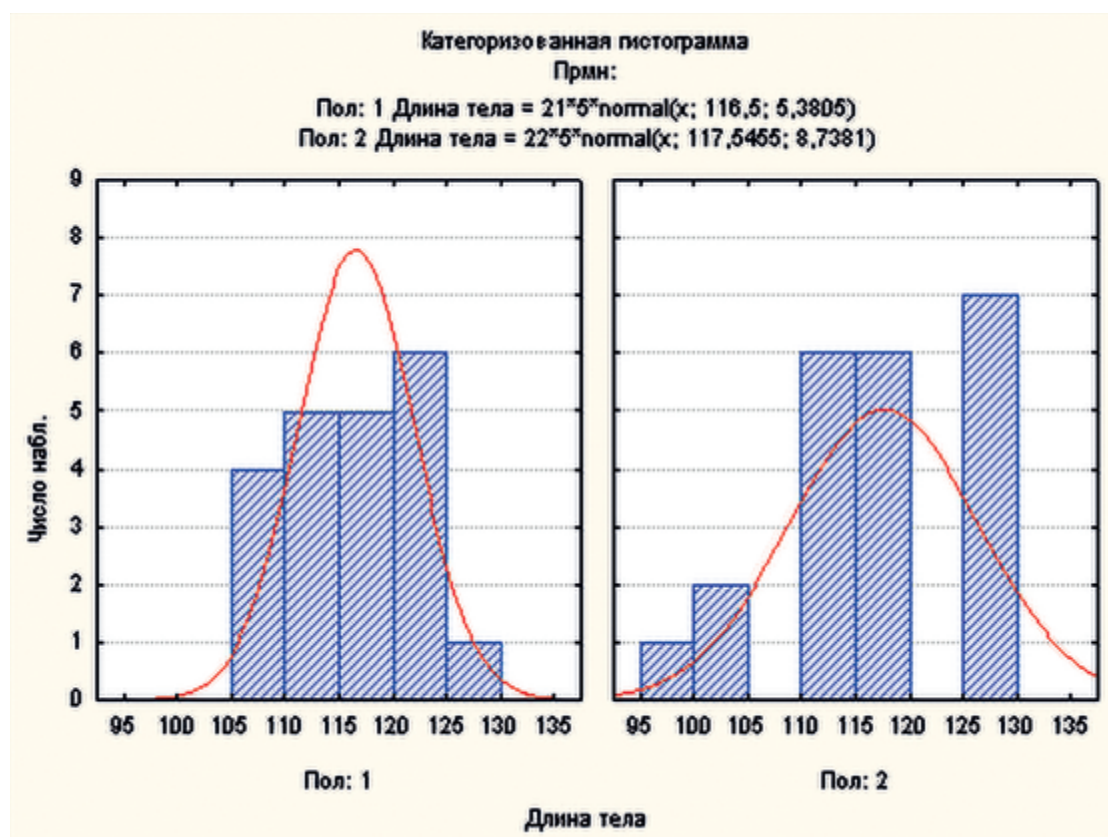


**Рисунок 5.**

Сравнительная характеристика длины тела девочек (1) и мальчиков (2) (критерий Манна-Уитни $U = 206,0$, $p = 0,54$): А – бокс-диаграмма или диаграмма размаха, Б – категоризованная гистограмма (анализ проведён в ППП Statistica).

Figure 5.

Comparison of the body length of girls (1) and boys (2) (Mann-Whitney U-test = 206.0, $p = 0.54$): A – box plot / span chart, B – histogram.



Итак, сформулируем выводы по рисунку 5:

- в изучаемой группе дошкольников длина тела колеблется от 100 до 130 см, причем у девочек колебания признака менее выражены – от 108 до 127 см (**рисунок 5А**), тогда как у мальчиков – от 100 до 130 см;
- максимальный рост отмечен в группе мальчиков (группа 2 по полу), впрочем, как и минимальный, – то есть размах значений среди мальчиков больше, чем среди девочек;
- в обеих группах распределение признака визуально отличается от нормального – от формы линии на рис. Б, – что предполагает использование непараметрических методов сравнения, например, критерия Манна-Уитни U;
- различия по длине тела между группами мальчиков и девочек статистически незначимы ($U = 206,0$, $p = 0,54$);
- имеется только статистически незначимая тенденция увеличения длины тела у мальчиков (**рисунок 5Б**), так как наибольшее число наблюдений среди девочек входит в интервал роста 120–125 см – 6 девочек, среди мальчиков – в интервал 125–130 см (7 мальчиков);
- распределение детей по интервалам категоризированной гистограммы на **рисунке 5Б** можно использовать для подбора учебной мебели.

Литература:

1. Боровиков В. *STATISTICA: искусство анализа данных на компьютере*. СПб.: Питер, 2001. 656 с.
2. Боровиков В.П. *Программа STATISTICA для студентов и инженеров*. 2-е изд. М.: Компьютер-Пресс, 2001. 301 с.
3. Гланц С. *Медико-биологическая статистика*. М.: Практика, 1998. 459 с.
4. Гудинова Ж.В., Жернакова Г.Н., Толькова Е.И. *Дружелюбная статистика: статистический анализ медицинских баз данных: пошаговые инструкции*. Вып. 1. Омск: Изд-во ОмГМА, 2014. 112 с.
5. ВОЗ. *Доклад о состоянии здравоохранения в Европе 2015. Целевые ориентиры и более широкая перспектива – новые рубежи в работе с фактическими данными*. Целевые ориентиры и более широкая перспектива – новые рубежи в работе с фактическими данными. 2017. 182 с.
6. *Статистика: учебник*. Под ред. проф. И.И. Елисевой. М.: ТК Велби, изд-во «Прспект», 2003. 448 с.
7. Зайцев В.М., Лифляндский В.Г., Маринкин В.И. *Прикладная медицинская статистика*. СПб.: Издательство «ФОЛИАНТ», 2003. 432 с.
8. Мерков А.М., Поляков Л.С. *Санитарная статистика: пособие для врачей*. Л.: Медицина, 1974. 384 с.
9. Microsoft. Справка и обучение по Excel [Электронный ресурс]. Ссылка активна на 29.01.2023. <https://support.microsoft.com/ru-RU>
10. Сепетлиев Д. *Статистические методы в научных медицинских исследованиях*. М.: Медицина, 1968. 419 с.
11. Петри А, Сэбин К. *Наглядная статистика в медицине*. Пер. с англ. под ред. В.П. Леонова. М.: ГЭОТАР-МЕД, 2003. 144 с.
12. Мангейм Д.Б., Рич Р.К. *Политология: Методы исследования*. М.: Изд-во «Весь Мир», 1997. 544 с. Ссылка активна на 29.01.2023. http://grachev62.narod.ru/Mr/Mr_14.html
13. Реброва О.Ю. *Статистический анализ медицинских данных. Применение пакета прикладных программ STATISTICA*. М.: МедиаСфера, 2006. 312 с.
14. *Социально-гигиенический мониторинг. Анализ медико-демографических и социально-экономических показателей на региональном уровне: Методические рекомендации* [Электронный ресурс]. М.: Федеральная служба по надзору в сфере защиты прав потребителей и благополучия человека, 2010. 53 с. Ссылка активна на 29.01.2023. <https://docs.cntd.ru/document/902236566>.
15. Гудинова Ж.В., Овчинникова Е.Л., Гегечкори И.В., Толькова Е.И., Жернакова Г.Н. *Социально-гигиенический мониторинг*. Омск: Изд-во ОмГМУ, 2016. 203 с.
16. Гудинова Ж.В., Овчинникова Е.Л., Гегечкори И.В., Толькова Е.И., Жернакова Г.Н. *Социально-гигиенический мониторинг: учебное пособие*. Омск: Изд-во ОмГМУ, 2017. 171 с.
17. Халафян А.А. *STATISTICA 6. Статистический анализ данных*. 3-е изд. М.: ООО «Бином-Пресс», 2007. 512 с.
18. Электронный учебник по статистике компании StatSoft [Электронный ресурс]. Ссылка активна на 29.01.2023. <http://www.statsoft.ru>.
19. Ланг Т.А., Сесик М. *Как описывать статистику в медицине: руководство для авторов, редакторов и рецензентов*. Пер. с англ. под ред. В.П. Леонова. М.: Практическая медицина, 2016. 470 с.

References:

1. Borovikov V. *STATISTICA: iskusstvo analiza dannykh na komp'yutere*. St. Petersburg: Peter; 2001. 656 p. (In Russ).
2. Borovikov VP. *Programma STATISTICA dlya studentov i inzhenerov*. 2nd ed Moscow: Komp'yuter-Press; 2001. 301 p. (In Russ).
3. Glantz S. *Mediko-biologicheskaya statistika*. Moscow: Praktika; 1998. 459 p. (In Russ).
4. Gudina ZhV, Zernakova GN, Tol'kova EI. *Druzhelyubnaya statistika: statisticheskiy analiz meditsinskikh baz dannykh: poshagovye instruktsii*. Publ 1. Omsk: Izd-vo OmGMA; 2014. 112 p. (In Russ).
5. WHO. *Report on the state of health in Europe 2015. Targets and a broader perspective – new frontiers in working with evidence*. 2017. 182 p. (In Russ).
6. Eliseeva II, editors. *Statistics: textbook*. Moscow: TK Velbi, izd-vo Prospekt; 2003. 448 p. (In Russ).
7. Zaytsev VM, Lifyandskiy VG, Marinkin VI. *Prikladnaya meditsinskaya statistika*. St. Petersburg: FOLIANT Publishing House, 2003; 432 p. (In Russ).
8. Merkov AM, Polyakov LS. *Sanitarnaya statistika: workbook*. Leningrad: Medicine; 1974. 384 p. (In Russ).
9. Microsoft. Excel Help [Electronic resource]. (Russia). Available at: <https://support.microsoft.com/ru-RU>. Accessed: January 29, 2023.
10. Sepetliev D. *Statisticheskie metody v nauchnykh meditsinskikh issledovaniyakh*. Moscow: Meditsina, 1968. 419 p. (In Russ).
11. Petri A. *Naglyadnaya statistika v meditsine*. Translated from englis. VP Leonova. Moscow: GEOTAR-MED, 2003. 144 p. (In Russ).
12. Mannheim JB, Rich RK. *Political science: Research methods*. Moscow: Publishing House The Whole World; 1997. 544 p. (In Russ). Available at: http://grachev62.narod.ru/Mr/Mr_14.html. Accessed: January 29, 2023.
13. Rebrova O.Y. *Statisticheskiy analiz meditsinskikh dannykh*. Primenenie paketa prikladnykh program. Moscow: Mediasphere; 2006. 312 p. (In Russ).
14. *Sotsial'no-gigienicheskiy monitoring. Analiz mediko-demograficheskikh i sotsial'no-ekonomicheskikh pokazateley na regional'nom urovne: Metodicheskie rekomendatsii* [Electronic

- resource]. Moscow: Federal'naya sluzhba po nadzoru v sfere zashchity prav potrebitel'nykh i blagopoluchiya cheloveka; 2010. 53 p. (In Russ). <https://docs.cntd.ru/document/902236566>. Accessed: January 29, 2023.
15. Gudina ZhV, Ovchinnikova EL, Gegechkori IV, Tol'kova EI, Zhernakova GN. *Sotsial'no-gigienicheskiy monitoring*. Omsk : Izd-vo OmGMA; 2016. 203 p. (In Russ).
 16. Gudina ZhV, Ovchinnikova EL, Gegechkori IV, Tol'kova EI, Zhernakova GN. *Sotsial'no-gigienicheskiy monitoring*. Omsk : Izd-vo OmGMA; 2017. 171 p. (In Russ).
 17. Khalafyan AA. *STATISTICA 6. Statisticheskiy analiz dannykh*. 3rd ed. Moscow : OOO Binom-Press; 2007. 512 p. (In Russ).
 18. Electronic textbook on statistics of the company StatSoft [Electronic resource]. (Russia). Available at: <http://www.statsoft.ru>. Accessed: January 29, 2023.
 19. Lang TA, Sesik M. *How to describe statistics in medicine: a guide for authors, editors and reviewers*. Translated from english VP Leonova. Moscow : Practical Medicine, 2016. 470 p. (In Russ).

Сведения об авторах

Гудина Жанна Владимировна, доктор медицинских наук, профессор, заведующая кафедрой общей гигиены, гигиены детей и подростков, ФГБОУ ВО «Омский государственный медицинский университет» Министерства здравоохранения Российской Федерации (644099, Россия, г. Омск, ул. Ленина, д. 12)

Вклад в статью: написание лекции.

ORCID: 0000-0001-6869-6057

Демакова Лидия Владимировна, кандидат медицинских наук, доцент кафедры общей гигиены, гигиены детей и подростков, ФГБОУ ВО «Омский государственный медицинский университет» Министерства здравоохранения Российской Федерации (644099, Россия, г. Омск, ул. Ленина, д. 12)

Вклад в статью: сбор материала, анализ литературных данных.

ORCID: 0000-0003-1742-7900

Статья поступила: 30.06.2022 г.

Принята в печать: 30.11.2022 г.

Контент доступен под лицензией CC BY 4.0.

Authors

Prof. Zhanna V. Gudina, MD, DSc, Professor, Head of the Department of General and Pediatric Hygiene, Omsk State Medical University (12, Lenina Street, Omsk, Russian Federation)

Contribution: wrote the manuscript.

ORCID: 0000-0001-6869-6057

Dr. Lidia V. Demakova, MD, PhD, Associate Professor, Department of General and Pediatric Hygiene, Omsk State Medical University (12, Lenina Street, Omsk, Russian Federation)

Contribution: collected the data; performed the data analysis.

ORCID: 0000-0003-1742-7900

Received: 30.06.2022

Accepted: 30.11.2022

Creative Commons Attribution CC BY 4.0.