

## Анализ парных выборок в биомедицинских исследованиях с использованием программного обеспечения SPSS и STATA: параметрические критерии

Гржибовский А.М.<sup>1-4</sup> PhD, заведующий центральной научной исследовательской лабораторией  
Унгурану Т.Н.<sup>1</sup> д.м.н., доцент кафедры гигиены и медицинской экологии  
Горбатова М.А.<sup>1</sup> к.м.н., научный сотрудник центральной научной исследовательской лаборатории

- 1 — ГБОУ ВПО «Северный государственный медицинский университет» Минздрава России  
163000, г. Архангельск, проспект Троицкий, д. 51
- 2 — Национальный Институт Общественного Здравоохранения  
Осло, Норвегия
- 3 — Северо-Восточный Федеральный Университет  
Якутск, Россия
- 4 — Международный Казахско-Турецкий Университет им. Х.А. Ясави  
Туркестан, Казахстан

Автор для корреспонденции: Гржибовский Андрей Мечиславович; e-mail: Andrej.Grijbovski@gmail.com

Финансирование. Исследование не имело спонсорской поддержки.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила: 13.06.2017.

*В настоящей статье рассмотрены способы сравнения количественных данных в парных группах с помощью параметрических критериев. Описано сравнение двух парных групп с помощью парного критерия Стьюдента, трех парных групп — с помощью однофакторного дисперсионного анализа повторных измерений. Представлены примеры расчетов вышеупомянутых критериев с использованием программного обеспечения SPSS и STATA. Даются рекомендации о том, как следует представлять результаты сравнения данных с помощью рассмотренных параметрических критериев в научных публикациях.*

**Ключевые слова:** статистический анализ, парные выборки, парный критерий Стьюдента, однофакторный дисперсионный анализ повторных измерений, SPSS, STATA.

В предыдущих выпусках практикума были представлены методы, применяемые для сравнения количественных данных в двух и более независимых выборках, при этом независимость сравниваемых групп являлось одним из условий применения этих методов. Настоящая статья знакомит читателей с параметрическими методами, которые используются для сравнения парных (зависимых) групп. В статье будут рассмотрены парный критерий Стьюдента (Paired-Samples T-test) для сравнения двух зависимых групп и однофакторный дисперсионный анализ повторных измерений (One-Way Repeated-Measures ANOVA), предназначенный для сравнения трех и более зависимых групп.

Под парными данными понимают значения связанных (зависимых) выборок. Парность изучаемых выборок определяется дизайном исследования. Примерами парных выборок могут служить ситуации, когда изучаемый признак измеряется у одних и тех же участников исследования в два разных момента

времени, например, до и после лечебного вмешательства, — исследования типа «до — после». Также парными данными будут являться, к примеру, результаты измерения одного и того же показателя у одной и той же группы пациентов: определение уровня холестерина в крови в нескольких лабораториях или измерение артериального давления двумя исследователями. Парными также считаются данные, полученные для двух групп в результате исследования типа «случай — контроль» с использованием метода подобранных пар [17]. В данном случае «контроли» не могут считаться полностью независимыми от «случаев», так как их намеренно подбирали, исходя из характеристик «случаев», например, пола, возраста, района проживания.

Статистическая обработка данных является завершающим этапом практически каждого эпидемиологического исследования, так как именно после выполнения статистического анализа полученных результатов принимается или отвергается исследовательская гипотеза. В связи

с этим авторы настоящей статьи рекомендуют читателям перед тем, как приступить к методам статистической обработки данных, ознакомиться с литературой по эпидемиологии и дизайнам исследований [2, 3, 16].

Принципиальное отличие сравнения связанных (парных) выборок от сравнения независимых выборок состоит в том, что в случае повторных наблюдений имеется связь между значениями измеряемого параметра для каждого наблюдения: значения параметра у одного и того же участника исследования в парных выборках будут ближе друг к другу, чем значения измеренного параметра у нескольких участников, и соответственно, дисперсия значений при повторных измерениях будет меньше [1, 4].

Параметрическим критерием для сравнения двух парных выборок является парный критерий Стьюдента. Парный критерий Стьюдента, предназначенный для сравнения средних величин двух связанных выборок, применяется только при соблюдении следующих условий:

1. Количественный тип данных, причем желательны непрерывные, а не дискретные данные.
2. Наличие не более двух связанных выборок.
3. Нормальное распределение разности между значениями изучаемого признака в группах.
4. Каждая из изучаемых пар должна быть независима от остальных пар наблюдений.

Для проверки соблюдения первых двух условий никаких математических манипуляций не требуется. Нормальность распределения разницы между значениями признака в связанных группах легко проверить путем создания новой переменной, каждое из значений которой будет представлять собой  $Y_2 - Y_1$ , где  $Y_2$  — значение переменной после вмешательства, а  $Y_1$  — значение этой переменной до вмешательства. Распределение новой переменной проверяется с помощью критериев Kolmogorov-Smirnov или Shapiro-Wilk, а также с помощью гистограмм и квантильных диаграмм. Независимость пар друг от друга также определяется дизайном исследования.

Если необходимо сравнить средние значения количественного признака для трех и более парных групп, то применяется дисперсионный анализ для повторных наблюдений (Repeated measures ANOVA) [4, 9]. Дисперсионный анализ для повторных измерений выделяет два источника вариабельности (уровни измерений и субъекты), в то время как при обычном дисперсионном анализе учитывается только вариабельность между субъектами.

Благодаря этому повышается чувствительность или статистическая мощность анализа. Вместо сравнения эффектов лечения в группе различных субъектов эффекты лечения оцениваются среди повторных измерений у одних и тех же субъектов. Каждый субъект, таким образом, служит собственным контролем. Следова-

тельно, данный тип дизайна исследования более чувствителен к разнице в значениях, то есть для того, чтобы отклонить ложную нулевую гипотезу, требуется меньшее количество участников исследования, чем при использовании дисперсионного анализа для независимых выборок при одних и тех же различиях между средними значениями зависимой переменной.

Нулевая гипотеза для дисперсионного анализа повторных наблюдений утверждает, что средние значения измерений не отличаются друг от друга, то есть разница средних значений для каждого повторного наблюдения равняется нулю. Соответственно альтернативная гипотеза заключается в том, что между средними значениями различия есть. Для применения дисперсионного анализа повторных наблюдений необходимо соблюдение ряда условий:

1. Условие нормального распределения изучаемой переменной.
2. Наличие трех или более связанных выборок.
3. Наблюдения должны быть независимыми (не выборки, а отдельные наблюдения).
4. Условие сферичности (приблизительное равенство дисперсий разности между всеми возможными уровнями пар значений). Необходимо отметить, что современные статистические программы позволяют использовать способы проведения дисперсионного анализа повторных измерений даже в случае несоблюдения условия сферичности.

В настоящей статье для приобретения читателями практических навыков проведения статистического сравнения количественных данных в парных выборках использованы примеры гипотетических исследований по изучению изменения содержания этанола в выдыхаемом воздухе (2 зависимых группы) и влиянию озонотерапии на уровень липидов у лиц, страдающих алкоголизмом (3 зависимых группы). С использованием программного обеспечения SPSS и STATA представлено сравнение двух зависимых групп с помощью парного критерия Стьюдента и сравнение трех зависимых групп — с помощью однофакторного дисперсионного анализа повторных измерений. Основную информацию по использованию SPSS можно получить в пособиях [8, 13, 18], информацию по использованию программы STATA можно прочитать в практикуме [5, 10—12, 14] и специализированной литературе [19—21].

### Сравнение двух зависимых групп с использованием программы SPSS 17.0

Расчет парного критерия Стьюдента в SPSS рассмотрен на примере данных гипотетического исследования, в котором изучалось содержание этанола в выдыхаемом воздухе у 71 испытуемого через 60 мин (переменная «Ethanol\_60») и 180 мин (переменная

«Ethanol\_180») после приема этанола в дозе 0,8 г/кг массы тела. Файл с примером «4\_Ethanol\_Lipid\_SPSS» можно загрузить с сайта журнала «Наркология».

В рассматриваемом примере нулевой будет являться гипотеза об отсутствии различий между средним содержанием этанола в выдыхаемом воздухе при измерении его концентрации через 60 мин и 180 мин после приема. Перед тем, как начать проверку гипотезы с помощью критерия Стьюдента для парных выборок, следует проверить, можно ли применять этот критерий в данной ситуации.

Исходя из дизайна исследования, условия наличия двух связанных выборок, количественных данных и независимости пар наблюдений соблюдаются. Для проверки соблюдения условия нормальности распределения разности между значениями изучаемого признака в группах необходимо создать новую переменную, обозначим ее «Ethanol\_dif», значения которой будут представлять собой разность между содержанием этанола через 180 мин и через 60 мин после его приема. Создается новая переменная путем открытия диалогового окна «Compute» (рис. 1) в меню «Transform».

Проверка распределения проводится с использованием графических методов или статистических критериев. Пошаговый алгоритм проверки распределения количественных данных на основе описательной статистики, графиков распределения и статистических критериев рассматривался в предыдущих выпусках практикума [5, 6], поэтому в настоящей статье представлен не будет.

Проверка распределения показала, что новая переменная подчиняется закону нормального распределения (табл. 1 и рис. 2), значит, среднее содержание этанола в выдыхаемом воздухе при измерении его концентрации через 60 мин и 180 мин после приема можно сравнивать с помощью парного критерия Стьюдента.

Для анализа связанных выборок данные в SPSS должны быть представлены в виде так называемого горизонтального файла, то есть файла, в котором для каждого из участников исследования значения изучаемой переменной регистрируются в одной строке. Для расчетов следует открыть диалоговое окно «Paired-Samples T Test», которое открывается при

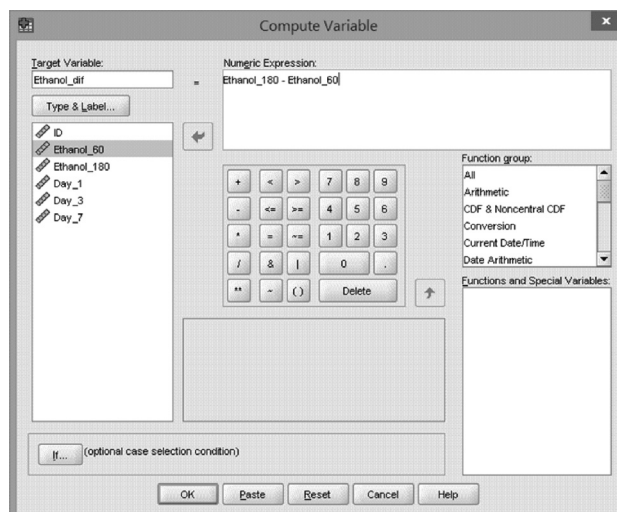


Рис. 1. Диалоговое окно «Compute Variable» для создания новых переменных с помощью математических вычислений программы SPSS 17.0.

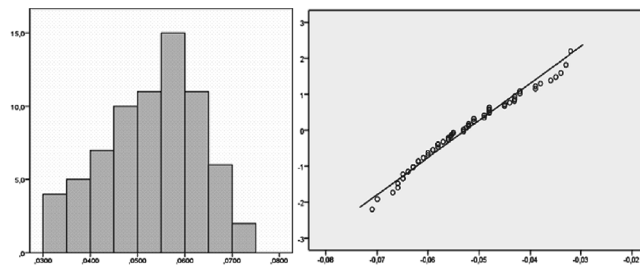


Рис. 2. Гистограмма (слева) и квантильная диаграмма (справа) распределения переменной «Ethanol\_dif».

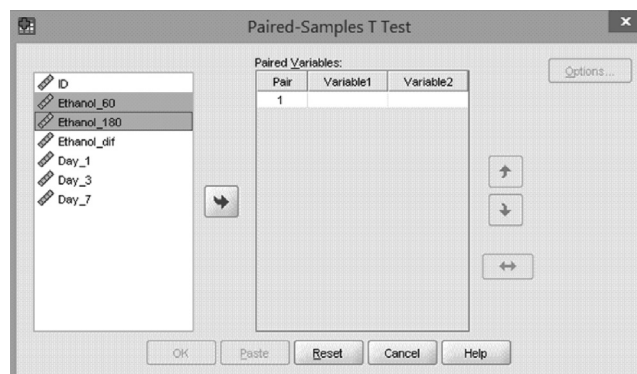


Рис. 3. Диалоговое окно «Paired-Samples T Test» программы SPSS 17.0.

Результаты проверки распределения переменной «Ethanol\_dif» с помощью статистических критериев

Таблица 1

**Tests of Normality**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Ethanol_dif	,076	71	,200 <sup>*</sup>	,973	71	,122

a. Lilliefors Significance Correction

\*. This is a lower bound of the true significance.

помощи меню «Analyze» → «Compare Means» → «Paired-Samples T Test» (рис. 3). В область «Paired Variables» путем нажатия на стрелку «►» переносится пара (или несколько пар) переменных, средние значения которых планируются сравнить. В данном примере это переменные «Ethanol\_60» и «Ethanol\_180», которые нужно отобразить и перенести одновременно. В меню «Options» можно изменить ширину доверительного интервала. По умолчанию рассчитывается 95% доверительный интервал.

Результаты применения парного критерия Стьюдента представлены в табл. 2 и 3. В табл. 2 показаны средние значения (Mean) содержания этанола через 60 мин и 180 мин после приема, а также число наблюдений (71), стандартные отклонения (Std. Deviation) и стандартные ошибки средних величин (Std. Error Mean).

В табл. 3 даны основные результаты сравнения групповых средних. Указано среднее значение разности содержания этанола (Mean), стандартное отклонение для разности значений (Std. Deviation), стандартная ошибка средней разности (Std. Error Mean), а также 95% доверительный интервал для средней арифметической (95% Confidence Interval of the Difference). В следующих столбцах представлено значение критерия Стьюдента для парных выборок (t), число степеней свободы (df) и уровень значимости (Sig. 2-tailed) для двустороннего теста, который позволяет отвергнуть нулевую гипотезу об отсутствии различий между средними значениями содержания этанола через 60 мин и 180 мин.

Представляя результаты применения парного критерия Стьюдента, рекомендуется указывать значение критерия, количество степеней свободы и достигнутый уровень значимости. Рассмотренный выше пример можно представить следующим образом: среднее значение содержания этанола через 60 мин ( $M = 0,15$  мг/л,  $SD = 0,018$ ) и через 180 мин ( $M = 0,09$  мг/л,  $SD = 0,021$ ) статистически различались ( $t_{70} = 45,7$ ,  $p < 0,001$ ).

**Сравнение трех зависимых групп с использованием программы SPSS 17.0**

Применение дисперсионного анализа повторных измерений в SPSS рассмотрено на примере данных гипотетического исследования, в котором изучалось влияние озонотерапии на содержание липидов у группы пациентов, страдающих алкоголизмом. Группу исследования составили 96 пациентов, которые в течение 7 дней дополнительно к стандартной терапии получали инфузию озонированного физиологического раствора хлорида натрия. Уровень липидов определяли через 1 сутки (переменная «Day\_1»), 3 суток (переменная «Day\_3») и 7 суток (переменная «Day\_7»).

Рассмотрим применение дисперсионного анализа повторных измерений для данного примера, сравнивая средние значения содержания липидов в три момента измерения. Будет проверяться нулевая гипотеза о том, что нет различий в среднем содержании липидов при трех измерениях.

Таблица 2

Описательная статистика для сравниваемых групп  
**Paired Samples Statistics**

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Ethanol_60	,15218	71	,018357	,002179
	Ethanol_180	,09954	71	,021918	,002601

Таблица 3

Результаты применения парного критерия Стьюдента для оценки различий между средними значениями связанных групп

**Paired Samples Test**

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower				Upper
Pair 1	Ethanol_60 - Ethanol_180	,052643	,009704	,001152	,050346	,054940	45,713	70	,000



Перед применением дисперсионного анализа повторных наблюдений необходимо проверить соблюдение условия нормальности распределения и условия сферичности с помощью критерия Mauchly. Проверка распределения показала, что распределение переменных «Day\_1», «Day\_3» и «Day\_7» подчиняется закону нормального распределения (табл. 4).

Для проведения дисперсионного анализа повторных наблюдений необходимо открыть диалоговое окно «Repeated Measures Define Factor(s)» (рис. 4) путем выбора в меню «Analyze» раздела «General Linear Model», в котором, в свою очередь, надо выбрать меню «Repeated Measures». В верхней строке (Within-Subject Factor Name) следует произвольно обозначить название изучаемого фактора (например, Lipidy, так как изучается содержание липидов). В строке «Number of Levels» нужно отметить количество изучаемых уровней, которое в данном случае равно количеству сравниваемых групп, то есть 3, затем нажать кнопку «Add», после чего SPSS позволит нажать на кнопку «Define» и открыть основное окно «Repeated Measures». Далее следует одновременно отметить все уровни изучаемой переменной (рис. 5) и переместить их в область «Within-Subjects Variables», заменив ими знаки вопроса в этой области.

Поскольку мы рассматриваем простейший вариант дисперсионного анализа повторных наблюдений, больше ничего в модель не вводится, а значит, нет смысла открывать меню «Model». В меню «Contrasts» по умолчанию установлено «Polynomial», что позволяет оценить тренд. При желании можно изменить тип плановых сравнений путем открытия окна «Contrasts» (рис. 6). При выборе «Deviation» каждая из групп будет сравниваться со средним значением всех наблюдений. При выборе «Simple» каждая из групп будет сравниваться либо с первой, либо с последней группой в зависимости от выбора исследователя (изменение первой группы на последнюю и наоборот осуществляется путем нажатия на «Change»). При выборе «Repeated» каждая последующая группа будет сравниваться с предыдущей, что удобно при изучении последовательных наблюдений.

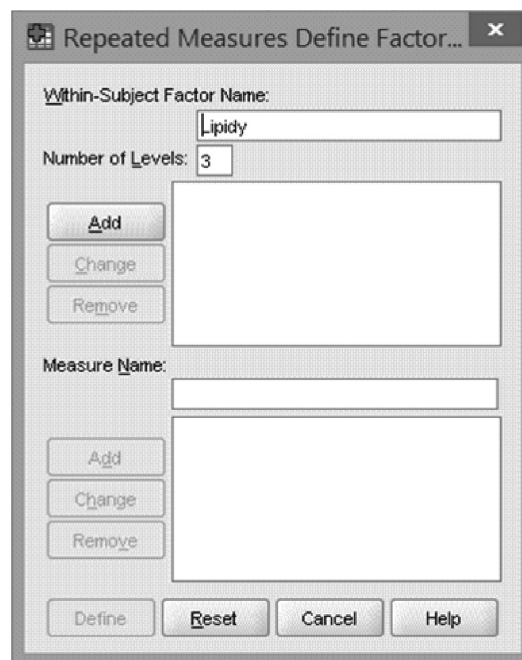


Рис. 4. Диалоговое окно «Repeated Measures Define Factor(s)» программы SPSS 17.0.

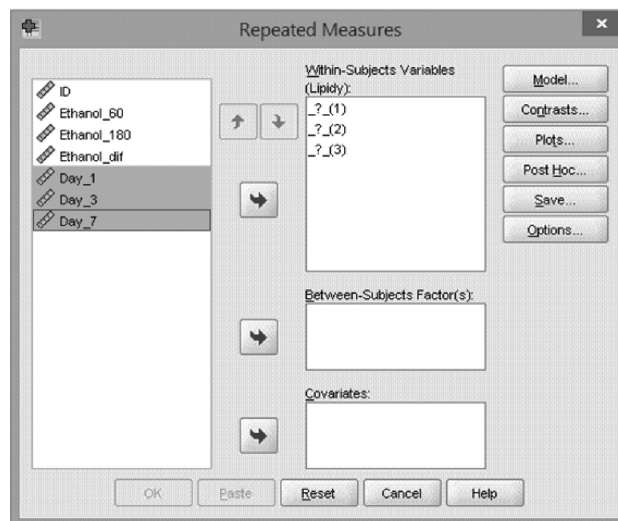


Рис. 5. Диалоговое окно «Repeated Measures» программы SPSS 17.0.

Таблица 4

Результаты проверки распределения переменных Day\_1», «Day\_3» и «Day\_7» с помощью статистических критериев

**Tests of Normality**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Day_1	,042	96	,200*	,991	96	,804
Day_3	,067	96	,200*	,980	96	,139
Day_7	,043	96	,200*	,989	96	,640

a. Lilliefors Significance Correction

\*. This is a lower bound of the true significance.

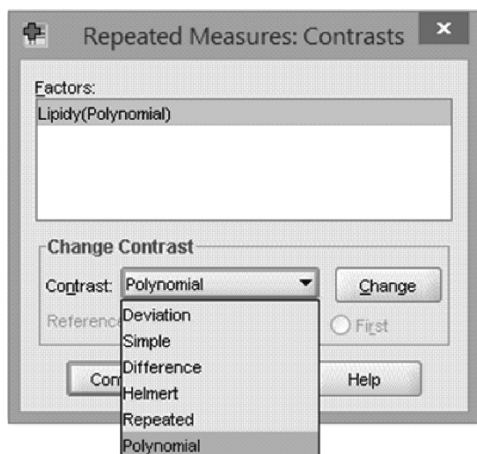


Рис. 6. Диалоговое окно «Repeated Measures: Contrasts» программы SPSS 17.0.

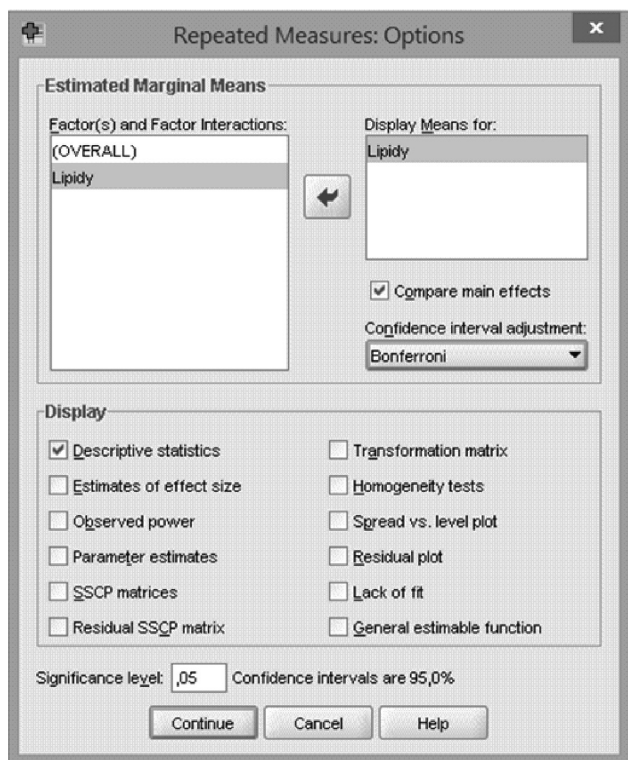


Рис. 7. Диалоговое окно «Repeated Measures: Options» программы SPSS 17.0.

Плановые сравнения типа «Helmert» обеспечивают сравнение каждой группы со средним значением всех оставшихся групп. Выбор типа плановых сравнений определяется исследователем, причем еще на этапе планирования исследования. В данном примере мы оставляем предложенный по умолчанию анализ тренда.

Поскольку в данном примере имеется всего один уровень для межгрупповых различий (вся выборка исследуется без разделения, например, по полу или возрасту), то выбор апостериорных тестов не имеет смысла и SPSS даже не предлагается. В меню «Options» можно выбрать апостериорные (post hoc) критерии для попарного сравнения изучаемых групп (рис. 7). Для этого нужно переместить изучаемый фактор (Lipidy) из области «Factor(s) and Factor Interactions» в область «Display Means for», отметить «Compare main effects» и выбрать один из предлагаемых критериев для апостериорных сравнений (LSD, Bonferroni и Sidak), из которых первый не рекомендуется, так как он не контролирует ошибку 1 типа. В данном примере выбрано сравнение по методу Bonferroni.

Учитывая наличие одной переменной, достаточно отметить только «Descriptive statistics» из всех предлагаемых программой функций и нажать на «Continue», после чего SPSS вернет «Repeated Measures» (рис. 5), в котором для запуска анализа необходимо нажать на «ОК».

Результаты дисперсионного анализа повторных наблюдений представлены в многочисленных таблицах, с каждой из которых попробуем разобраться отдельно. В табл. 5 представлены среднее арифметическое и стандартное отклонение для каждой из трех изучаемых переменных. Из таблицы видно, что среднее содержание липидов снижается с течением времени, однако пока невозможно сказать, являются ли различия статистически значимыми.

Таблица, обозначенная «Mauchly's Test of Sphericity» (табл. 6), показывает результат проверки соблюдения одного из необходимых условий — сферичности. В столбце «Sig.» представлен уровень значимости для критерия Mauchly, который показывает, что нулевую гипотезу о наличии сферичности можно

Таблица 5

Описательная статистика

**Descriptive Statistics**

	Mean	Std. Deviation	N
Day_1	123,0802	2,64518	96
Day_3	102,3781	3,33669	96
Day_7	101,3458	3,24391	96

отвергнуть, значит, одно из необходимых условий не соблюдается. Что делать в этом случае? В программе предусмотрены несколько видов коррекции для ситуаций отклонения от сферичности. Для оценки сферичности программа представляет значение Epsilon, которое может принимать значение от  $1 / (n - 1)$ , где  $n$  — количество уровней, до 1, то есть в данном случае Epsilon может принимать значение от 0,5 (Lower bound) до 1. Как видно из таблицы, значение Epsilon равно 0,649, что гораздо ближе к 0,5, чем к 1, что говорит о значительном отклонении от сферичности.

В таблице «Test of Within-Subjects Effects» (табл. 7) представлены основные результаты анализа. Однако помимо знакомого отношения F в таблице их целых 4. Какой из них выбрать? Перед тем как определиться с выбором, рассмотрим, что показано в этой таблице.

Во втором столбце представлено значение вариативности содержания липидов, обусловленное моделью (Lipidy Type III Sum of Squares, в данном слу-

чае, фактом измерения в различные промежутки времени), и остаточная вариативность, которая не объясняется моделью (Error Type III Sum of Squares). После деления этого значения на число степеней свободы (df) рассчитывается систематическая дисперсия (Lipidy Mean Square) и остаточная дисперсия (Error Mean Square). Отношение F получается в результате деления первого на второе. Уровень значимости для отношения F представлен в столбце «Sig.». В первой строке показано значение критерия F для ситуации, при которой соблюдается условие сферичности (Sphericity Assumed), а значит, для данного примера не подходит. В двух последующих строках видим результаты с коррекциями на отклонения от сферичности по Greenhouse-Geisser и Huynh-Feldt. Который из них выбрать? Рекомендуется при значении Epsilon  $< 0,75$  применять коррекцию по Greenhouse-Geisser, а при Epsilon  $> 0,75$  — коррекцию по Huynh-Feldt. Поскольку для рассматриваемого примера Epsilon  $< 0,75$ , то результат дисперсионного анализа смотрим

Таблица 6

Результат проверки сферичности с помощью критерия Mauchly

### Mauchly's Test of Sphericity<sup>b</sup>

Measure: MEASURE 1							
Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon <sup>a</sup>		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Lipidy	,459	73,276	2	,000	,649	,654	,500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept  
Within Subjects Design: Lipidy

Таблица 7

Таблица результатов дисперсионного анализа (F-test) для связанных выборок

### Tests of Within-Subjects Effects

Measure: MEASURE 1						
Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Lipidy	Sphericity Assumed	28864,798	2	14432,399	3252,460	,000
	Greenhouse-Geisser	28864,798	1,298	22245,773	3252,460	,000
	Huynh-Feldt	28864,798	1,308	22067,691	3252,460	,000
	Lower-bound	28864,798	1,000	28864,798	3252,460	,000
Error(Lipidy)	Sphericity Assumed	843,102	190	4,437		
	Greenhouse-Geisser	843,102	123,266	6,840		
	Huynh-Feldt	843,102	124,261	6,785		
	Lower-bound	843,102	95,000	8,875		

по второй строке, согласно ему можно отвергнуть нулевую гипотезу о сходстве дисперсий, а значит, и средних значений содержания липидов между группами. Это означает, что, по результатам проведенного дисперсионного анализа повторных наблюдений с поправкой Greenhouse-Geisser, есть статистически значимые различия в содержании липидов при использовании озонотерапии через 1 сутки, 3 суток и 7 суток.

При представлении результатов дисперсионного анализа парных наблюдений рекомендуется указывать значение отношения F, количество степеней свободы и достигнутый уровень значимости. В данном случае  $F_{2, 123} = 3252,4$ ,  $p < 0,001$ . Кроме того, учитывая, что количество степеней свободы кажется странным при наличии трех групп по 96 наблюдений

в каждой, логичным будет сообщить, что число степеней свободы было изменено в результате проведения коррекции по Greenhouse-Geisser по причине несоблюдения условия сферичности.

В табл. 8 представлены результаты оценки линейного и квадратичного трендов — оба являются статистически значимыми (значения в столбце «Sig.» меньше 0,05).

Апостериорные сравнения с поправкой Bonferroni представлены в таблице «Pairwise Comparisons» (табл. 9). Результаты указывают на существование статистически значимых различий ( $p < 0,001$ ) между группами 1 и 2, 2 и 3, 1 и 3, то есть установлены попарные различия в среднем содержании уровней липидов, измеренных через 1 сутки, 3 суток и 7 суток после применения озонотерапии.

Таблица 8

Плановые сравнения (анализ трендов)

**Tests of Within-Subjects Contrasts**

Measure:MEASURE\_1

Source	Lipidy	Type III Sum of Squares	df	Mean Square	F	Sig.
Lipidy	Linear	22674,387	1	22674,387	3765,336	,000
	Quadratic	6190,411	1	6190,411	2169,880	,000
Error(Lipidy)	Linear	572,078	95	6,022		
	Quadratic	271,024	95	2,853		

Таблица 9

Результаты апостериорных сравнений

**Pairwise Comparisons**

Measure:MEASURE\_1

(I) Lipidy	(J) Lipidy	Mean Difference (I-J)	Std. Error	Sig. <sup>a</sup>	95% Confidence Interval for Difference <sup>a</sup>	
					Lower Bound	Upper Bound
1	2	20,702 <sup>*</sup>	,357	,000	19,832	21,572
	3	21,734 <sup>*</sup>	,354	,000	20,871	22,598
2	1	-20,702 <sup>*</sup>	,357	,000	-21,572	-19,832
	3	1,032 <sup>*</sup>	,156	,000	,651	1,413
3	1	-21,734 <sup>*</sup>	,354	,000	-22,598	-20,871
	2	-1,032 <sup>*</sup>	,156	,000	-1,413	-,651

Based on estimated marginal means

\*. The mean difference is significant at the ,05 level.

a. Adjustment for multiple comparisons: Bonferroni.



### Сравнение двух зависимых групп с использованием программы STATA 14.0

Расчет парного критерия Стьюдента в STATA рассмотрен на примере данных гипотетического исследования, в котором изучалось содержание этанола в выдыхаемом воздухе у 71 испытуемого через 60 мин (переменная «Ethanol\_60») и 180 мин (переменная «Ethanol\_180») после приема этанола в дозе 0,8 г/кг массы тела.

Следует проверить соблюдение условия нормальности распределения разности между значениями изучаемого признака в группах. Для того чтобы рассчитать разность между парными наблюдениями, создадим новую переменную Ethanol\_dif в STATA. Для создания новой переменной, получаемой путем арифметического вычисления, необходимо открыть Data > Create or change data > Create new variable. В появившемся окне «generate — Create a new variable» в поле «Variable name» вписать название создаваемой переменной, например «Ethanol\_dif». Затем нажать на кнопку Create и в появившемся окне «Expression builder» в поле Category выбрать Variables. Выбирая нужные переменные из списка (дважды кликнув на переменную) и арифметические действия, отображаемые в окне справа, вводим формулу расчета новой переменной, например Ethanol\_180 — Ethanol\_60 (рис. 8). После нажатия на «ОК» в базе данных появится новый столбец с переменной «Ethanol\_dif». Создать новую переменную можно с помощью команды «generate». Для этого в нижней части главного диалогового окна STATA в поле Commands необходимо ввести: `. generate Ethanol_dif = Ethanol_180 — Ethanol_60`.

Проверка распределения проводится с использованием описательной статистики, графических методов и статистических критериев. Алгоритм построения гистограммы, квантильной диаграммы и расчета статистических тестов на нормальное распределение в STATA подробно был изложен в предыдущих но-



Рис. 8. Диалоговое окно для создания новой переменной путем вычисления в STATA 14.0.

мерах журнала [5, 6] и в практикуме [10], поэтому в настоящей статье не описывается.

Результаты проверки распределения переменной «Ethanol\_dif» с помощью статистического критерия Shapiro-Wilk представлены в табл. 10. Достигнутый уровень статистической значимости (в данном случае обозначается как Prob>z) составил 0,168, следовательно, распределение переменной «Ethanol\_dif» не отличается от нормального распределения.

Для сравнения средних значений с помощью парного критерия Стьюдента необходимо в меню «Statistics» выбрать «Summaries, tables, and tests», далее выбрать «Classical tests of hypotheses», а затем «t test (mean-comparison test)». В верхней части открывшегося окна «t test (mean-comparison test)» сле-

Таблица 10

Результаты проверки распределения переменной «Ethanol\_dif» с помощью статистического критерия Shapiro-Wilk (команда: `. swilk`)

Variable	Obs	W	V	z	Prob>z
Ethanol_dif	71	0.97502	1.556	0.962	0.16814

дует выбрать «Paired», в нижней части этого окна в поле «First variable» нужно перенести тестируемую переменную «Ethanol\_60», а в поле «Second variable» перенести переменную «Ethanol\_180» (рис. 9). В поле «Confidence level» можно изменить ширину доверительного интервала. По умолчанию рассчитывается 95% доверительный интервал. Для расчета парного критерия Стьюдента также можно воспользоваться командой: `. ttest Ethanol_60 == Ethanol_180`.

В табл. 11 представлено количество наблюдений в каждой группе (Obs), средние арифметические (Mean), ошибки среднего (Std. Err.), стандартные отклонения (Std. Dev.), 95% доверительные интер-

валы (95% Conf. Interval). Внизу таблицы указано значение t-критерия ( $t = 45,70$ ), количество степеней свободы ( $df = 70$ ), а также достигнутые уровни значимости для односторонней и двусторонней статистических гипотез. Значение достигнутого уровня статистической значимости для двустороннего критерия составило  $p < 0,001$ , что позволяет отклонить нулевую гипотезу об отсутствии различий между группами, следовательно, средние значения содержания этанола через 60 мин и 180 мин на популяционном уровне различаются.

### Сравнение трех зависимых групп с использованием программы STATA 14.0

Рассмотрим применение однофакторного дисперсионного анализа повторных измерений в STATA на примере данных гипотетического исследования, в котором изучалось влияние озонотерапии на содержание липидов у группы пациентов, страдающих алкоголизмом. Группу исследования составили 96 пациентов, которые в течение 7 дней дополнительно к стандартной терапии получали инфузию озонированного физиологического раствора хлорида натрия. Уровень липидов определяли через 1 сутки (переменная «Day1»), 3 суток (переменная «Day3») и 7 суток (переменная «Day7»). В рассматриваемом примере нулевой будет являться гипотеза об отсутствии различий между средним содержанием липидов при трех измерениях.

Перед применением дисперсионного анализа повторных наблюдений необходимо проверить соблюдение условия нормальности распределения и условия сферичности с помощью критерия Mauchly. Провер-

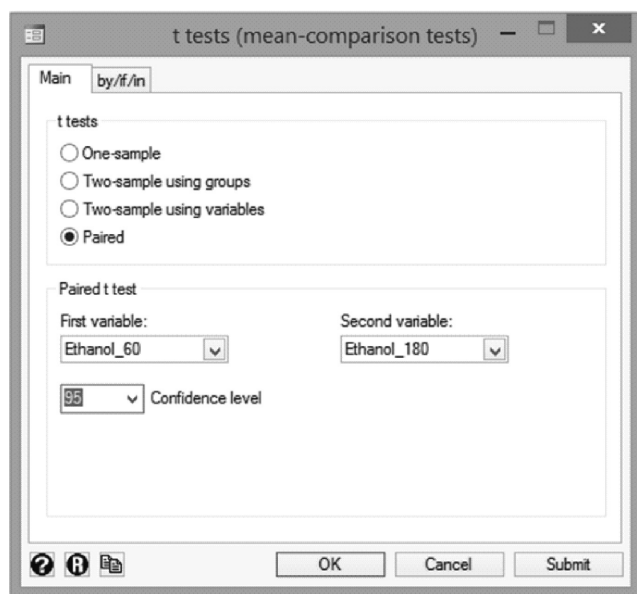


Рис. 9. Диалоговое окно «t test (mean-comparison test)» программы STATA 14.0.

Результаты применения парного критерия Стьюдента для оценки различий средних величин по данным программы STATA 14.0

Таблица 11

```
. ttest Ethanol_60 == Ethanol_180
```

Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Ethan~60	71	.1521831	.0021787	.0183578	.1478379	.1565283
Etha~180	71	.0995493	.0026013	.0219185	.0943613	.1047373
diff	71	.0526338	.0011517	.0097045	.0503368	.0549308

```

      mean(diff) = mean(Ethanol_60 - Ethanol_180)          t = 45.7002
Ho: mean(diff) = 0                                         degrees of freedom = 70

Ha: mean(diff) < 0      Ha: mean(diff) != 0      Ha: mean(diff) > 0
Pr(T < t) = 1.0000      Pr(|T| > |t|) = 0.0000      Pr(T > t) = 0.0000

```

ка распределения с использованием критерия Shapiro-Wilk показала, что распределение переменных «Day1», «Day3» и «Day7» подчиняется закону нормального распределения (табл. 12).

Для проведения дисперсионного анализа повторных наблюдений в STATA необходимо изменить формат данных на вертикальный. В STATA данные могут содержаться в двух форматах — горизонтальном (wide) и вертикальном (long). Для перекодировки формата данных можно воспользоваться командой: `. reshape long Day, i(ID) j(time)`. В случае правильного применения команды в рабочем окне появляется таблица 13.

Для проведения дисперсионного анализа повторных наблюдений необходимо в меню «Statistics» выбрать «Linear models and related», далее выбрать «ANOVA/MANOVA» и затем «Analysis of variance and covariance». В открывшемся диалоговом окне «anova — Analysis of variance and covariance» (рис. 10) в поле «Dependent variable» выбрать зависимую переменную «Day», в поле «Model» — ID time, отметить  Repeated-measured variables и в этом поле выбрать «time» и нажать ОК. Дисперсионный

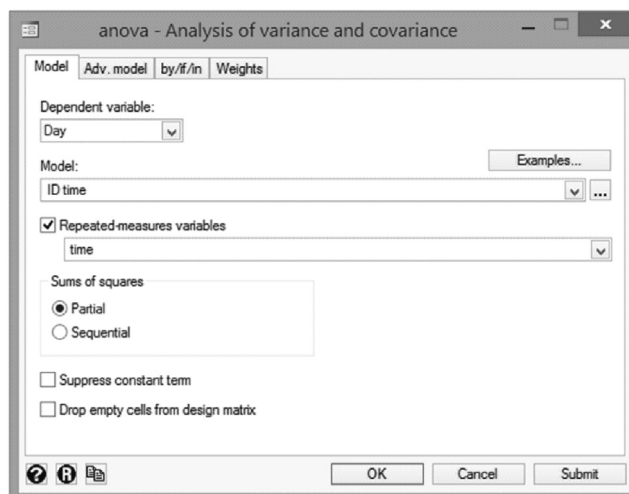


Рис. 10. Диалоговое окно «anova — Analysis of variance and covariance» программы STATA 14.0.

анализ повторных наблюдений можно выполнить используя команду: `. anova Day ID time, repeated(time)`.

Опция `repeated()` специфицирует переменную, указывающую на уровни измерения переменной

Таблица 12

Результаты проверки распределения переменных Day1, «Day3» и «Day7» с помощью критерия Shapiro-Wilk

```
. swilk Day1 Day3 Day7
```

Shapiro-Wilk W test for normal data						
Variable	Obs	W	V	z	Prob>z	
Day1	96	0.99246	0.602	-1.124	0.86943	
Day3	96	0.97974	1.617	1.063	0.14384	
Day7	96	0.98972	0.821	-0.438	0.66914	

Таблица 13

Отчет STATA об изменении формата данных

```
. . reshape long Day, i(ID) j(time)
(note: j = 1 3 7)
```

Data	wide	->	long
Number of obs.	96	->	288
Number of variables	6	->	5
j variable (3 values)		->	time
xij variables:	Day1 Day3 Day7	->	Day

(в примере после преобразования данных переменная time принимает значения 1, 3 и 7, обозначая три измерения содержания липидов — через 1 сутки, через 3 суток и через 7 суток. В табл. 14 представлены результаты дисперсионного анализа. Во втором столбце приведено значение вариабельности значений липидов (Partial SS time) и остаточная вариабельность, которая не объясняется моделью (Residual). Путем деления на число степеней свободы (df) рассчитывается систематическая дисперсия (MS time) и остаточная дисперсия (MS residual). Отношение F получается в результате деления этих двух значений.

Уровень значимости для отношения F представлен в столбце Prob > F. В строке time показан достигнутый уровень значимости для ситуации, при которой соблюдается условие сферичности. Это же значение находим в следующей табл. 15 — Regular Prob > F. Однако для того чтобы решить, можно ли использовать данное значение уровня значимости, необходимо

оценить, выполняется ли для данных условие сферичности.

Как было выше указано, сферичность можно оценить с помощью значения Epsilon, которое может принимать значение от  $1/(n-1)$ , где n — количество уровней, до 1, то есть в данном случае Epsilon может принимать значение от 0,5 (Lower bound) до 1. Как видно из табл. 15, значение Epsilon равно 0,648, что гораздо ближе к 0,5, чем к 1, что говорит о значительном отклонении от сферичности. Таким образом, в примере условие сферичности не соблюдается.

В столбцах таблицы 15 Prob > F H-F, G-G и Box находятся скорректированные на отсутствие сферичности уровни значимости. При значении Epsilon < 0,75 рекомендуется применять коррекцию по Greenhouse-Geisser, а при Epsilon > 0,75 — коррекцию по Huynh-Feldt. Поскольку для рассматриваемого примера Epsilon < 0,75, то результат дисперсионного анализа следует смотреть в столбце G-G.

Таблица 14

Результаты дисперсионного анализа (F-test) для связанных выборок при условии соблюдения условия сферичности

```

. anova Day ID time, repeated(time)

```

	Number of obs =	288	R-squared =	0.9733
	Root MSE =	2.10651	Adj R-squared =	0.9597

Source	Partial SS	df	MS	F	Prob>F
Model	30743.771	97	316.94609	71.43	0.0000
ID	1878.9728	95	19.778661	4.46	0.0000
time	28864.798	2	14432.399	3252.46	0.0000
Residual	843.10201	190	4.437379		
Total	31586.873	287	110.05879		

Таблица 15

Результаты дисперсионного анализа (F-test) для связанных выборок, скорректированные на отсутствие выполнения условия сферичности

```

Between-subjects error term: ID
Levels: 96 (95 df)
Lowest b.s.e. variable: ID

Repeated variable: time
Huynh-Feldt epsilon = 0.6540
Greenhouse-Geisser epsilon = 0.6488
Box's conservative epsilon = 0.5000

```

Source	df	F	Prob > F			
			Regular	H-F	G-G	Box
time	2	3252.46	0.0000	0.0000	0.0000	0.0000
Residual	190					



Согласно данному результату нулевую гипотезу о сходстве дисперсий ( $p < 0,001$ ) можно отвергнуть. Это означает, что, по результатам проведенного дисперсионного анализа результатов наблюдений с поправкой Greenhouse-Geisser, есть статистически значимые различия в содержании липидов через 1 сутки, 3 суток и 7 суток при использовании озонотерапии.

При представлении результатов дисперсионного анализа парных наблюдений рекомендуется указывать значение отношения  $F$ , число степеней свободы и достигнутый уровень значимости. В данном случае  $F_{2,190} = 3252,46$ ,  $p < 0,001$ . Подробнее о представлении результатов дисперсионного анализа повторных наблюдений, а также других одномерных методов статистического анализа можно ознакомиться в [7, 15].

### Заключение

В настоящей статье рассмотрены способы сравнения количественных данных в парных группах с помощью параметрических критериев. Описано сравнение двух парных групп с помощью парного критерия Стьюдента, трех парных групп — с помощью однофакторного дисперсионного анализа повторных измерений. Напомним, что данные методы являются параметрическими, а потому могут применяться в случае, если соблюдается условие нормальности распределения данных. В следующем выпуске практикума будет рассмотрено сравнение парных групп с помощью непараметрических критериев.

### Список литературы

1. Банержи А. Медицинская статистика понятным языком: вводный курс. М.: Практическая медицина, 2007. 287 с.
2. Власов В. В. Эпидемиология: учебное пособие для вузов. М.: ГЭОТАР-МЕД, 2004. 464 с.
3. Гланц С. Медико-биологическая статистика. М.: Практика, 1998. 459 с.
4. Гржибовский А.М. Одномерный анализ повторных измерений. *Экология человека* 2008; (4): 51-60.
5. Гржибовский А.М., Унгурияну Т.Н., Горбатова М.А. Описательная статистика с использованием пакетов статистических программ SPSS и STATA. *Наркология* 2017; (4): 36-51.
6. Гржибовский А.М., Унгурияну Т.Н., Горбатова М.А. Анализ непарных выборок в биомедицинских исследованиях с использованием программного обеспечения SPSS и STATA: параметрические критерии. *Наркология* 2017; (5): 18-36.
7. Ланг Т.А., Сесик М. Как описывать статистику в медицине. Аннотированное руководство для авторов, редакторов и рецензентов. Перевод с англ. под ред. В.П. Леонова. М.: Практическая медицина, 2011. 480 с.
8. Наследов А.Д. SPSS: Компьютерный анализ данных в психологии и социальных науках. СПб.: Питер, 2007. 416 с.
9. Петри А., Сэбин К. Наглядная медицинская статистика. М.: ГЭОТАР-Медиа, 2009. 168 с.
10. Субботина А.В., Гржибовский А.М. Описательная статистика и проверка нормальности распределения количественных данных. *Экология человека* 2014; (2): 60-63.
11. Субботина А.В., Гржибовский А.М. Одномерный дисперсионный анализ повторных наблюдений с использованием па-

кета статистических программ STATA. *Экология человека* 2014; (11): 61-64.

12. Субботина А.В., Гржибовский А.М. Сравнение количественных данных в двух парных выборках с использованием пакета статистических программ STATA. *Экология человека* 2014; (10): 61-64.
13. Таганов Д. SPSS: Статистический анализ в маркетинговых исследованиях. СПб.: Питер, 2005. 192 с.
14. Унгурияну Т.Н., Гржибовский А.М. Программное обеспечение для статистической обработки данных STATA: введение. *Экология человека* 2014; (1): 60-63.
15. Унгурияну Т.Н., Гржибовский А.М. Краткие рекомендации по описанию, статистическому анализу и представлению данных в научных публикациях. *Экология человека* 2011; (5): 55-60.
16. Флетчер Р., Флетчер С., Вагнер Э. Клиническая эпидемиология: Основы доказательной медицины. М.: МедиаСфера, 1998. 345 с.
17. Холматова К.К., Гржибовский А.М. Применение исследований «случай-контроль» в медицине и общественном здравоохранении. *Экология человека* 2016; (8): 53-60.
18. Field A. *Discovering statistics using SPSS*. London, 2005. 781 p.
19. Hamilton L. C. *Statistics with STATA: Updated for Version 10*. Brooks/Cole, Cengage Learning, 2009. 491 p.
20. Kohler U. Kreute F. *Data Analysis Using Stata*. USA, Texas: Stata Press, 2005. 378 p.
21. Rabe-Hesketh, S., Everitt, Brian. *A Handbook of Statistical Analyses Using Stata*. New York: Chapman & Hall, 2007. 352 p.

### References

1. Banerjee A. *Meditsinskaya statistika ponyatnym yazykom: vvodnyi kurs* [Medical statistics made clear]. Moscow: Prakticheskaya meditsina, 2007. 287 p. (In Russ.)
2. Vlasov V.V. *Epidemiologiya: uchebnoe posobie dlya vuzov* [Epidemiology: tutorial for high schools]. Moscow: GEOTAR-MED, 2004. 464 p. (In Russ.)
3. Glantz S. *Mediko-biologicheskaya statistika* [Biostatistica]. Moscow: Praktika, 1998. 459 p. (In Russ.)
4. Grijbovski A.M. *Odnomernyj analiz povtornykh izmerenij*. [Univariate analysis of repeated measurements]. *Ekologiya cheloveka* [Human Ecology] 2008; (4): 51-60. (In Russ.)
5. Grijbovski A.M., Unguryanu T.N., Gorbatova M.A. *Opisatel'naya statistika s ispol'zovaniem paketov statisticheskikh program SPSS i STATA*. [Descriptive statistics with using SPSS and STATA]. *Narkologiya* [Narkology] 2017; (4): 36-51. (In Russ.)
6. Grijbovski A.M., Unguryanu T.N., Gorbatova M.A. *Analiz neparnykh vyborok v biomeditsinskikh issledovaniyakh s ispol'zovaniem programmnogo obespecheniya SPSS i STATA: parametricheskie kriterii* [Analysis of non-paired samples in biomedical studies using SPSS and STATA software: parametric tests]. *Narkologiya* [Narkology] 2017; (5): 18-36. (In Russ.)
7. Nasledov A. D. *SPSS: Komp'yuternyy analiz dannykh v psikhologii i sotsial'nykh naukakh*. [SPSS Computer data analysis in psychology and social sciences]. Saint Petersburg: Piter, 2007. 416 p. (In Russ.)
8. Lang T.A., Secic M. *Kak opisivat' statistiku v medicine*. *Annotirovannoe rukovodstvo dlja avtorov, redaktorov i recenentov*. [How to report statistics in Medicine. Annotated guidelines for authors, editors, and reviewers]. Moscow: Practical Medicine, 2011. 480 p. (in Russ.)
9. Petri A., Sebin K. *Naglyadnaya meditsinskaya statistika*. [Medical statistics at glance]. Moscow: GEOTAR-Media, 2009. 168 p. (In Russ.)
10. Subbotina A. V., Grijbovski A. M. *Opisatel'naya statistika i proverka normal'nosti raspredeleniya kolichestvennykh dannykh*. [Descriptive statistics and normality testing for quantitative data].

*Ekologiya cheloveka [Human Ecology]* 2014; (2): 60-63. (In Russ.)

11. Subbotina A.V., Grjibovski A.M. Odnomernyj dispersionnyj analiz povtornyh nabljudenij s ispol'zovaniem paketa statisticheskikh programm STATA. [Analysis of variance for repeated measurements using STATA]. *Ekologiya cheloveka [Human Ecology]* 2014; (11): 61 — 64. (In Russ.)

12. Subbotina A.V., Grjibovski A.M. Sravnenie kolichestvennyh dannyh v dvuh parnyh vyborkah s ispol'zovaniem paketa statisticheskikh programm STATA. [Comparison of quantitative data in two paired samples using STATA software]. *Ekologiya cheloveka [Human Ecology]* 2014; (10): 61 — 64. (In Russ.)

13. Taganov D. SPSS: Statisticheskij analiz v marketingovykh issledovaniyakh. [SPSS: Statistical analysis in marketing research]. Saint Petersburg: Piter, 2005. 192 p. (In Russ.)

14. Unguryanu T.N., Grjibovski A.M. Programmnoe obecpechenie dlya statisticheskoy obrabotki dannyh STATA: vvedenie. [Introduction to STATA — Software for statistical data analysis]. *Ekologiya cheloveka [Human Ecology]* 2014; (1): 60-63. (In Russ.)

15. Unguryanu T.N., Grjibovski A.M. Kratkie rekomendacii po opisaniyu, statisticheskomu analizu i predstavleniju dannyh v nauchnyh publikacijah. [Brief recommendations on description, analysis and presentation of data in scientific papers]. *Ekologiya cheloveka [Human Ecology]* 2011; (5): 55 — 66. (In Russ.)

16. Fletcher R., Fletcher S., Vagner E. Klinicheskaya epidemiologiya: Osnovy dokazatel'noy meditsiny. [Clinical Epidemiology: the essentials]. Moscow: MediaSfera, 1998. 345 p. (In Russ.)

17. Kholmatova K.K., Grjibovski A.M. Primenenie issledovaniy «sluchay-kontrol'» v meditsine i obshchestvennom zdravookhraneni. [Case-control studies in medicine and public health]. *Ekologiya cheloveka [Human Ecology]* 2016; (8): 53 — 60. (In Russ.)

18. Field A. Discovering statistics using SPSS. London, 2005. 781 p.

19. Hamilton L.C. Statistics with STATA: Updated for Version 10. Brooks/Cole, Cengage Learning, 2009. 491 p.

20. Kohler U. Kreute F. Data Analysis Using Stata. USA, Texas: Stata Press, 2005. 378 p.

21. Rabe-Hesketh, S., Everitt, Brian. A Handbook of Statistical Analyses Using Stata. New York: Chapman & Hall, 2007. 352 p.

## ANALYSIS OF PAIRED SAMPLES IN BIOMEDICAL STUDIES USING SPSS AND STATA SOFTWARE: PARAMETRIC TESTS

**Grjibovski A.M.<sup>1-4</sup>, Unguryanu T.N.<sup>1,5</sup>, Gorbatova M.A.<sup>1</sup>**

1 — Northern State Medical University  
Arkhangelsk, Russia

2 — Norwegian Institute of Public Health  
Oslo, Norway

3 — North-Eastern Federal University  
Yakutsk, Russia

4 — International Kazakh-Turkish University  
Turkestan, Kazakhstan

5 — Arctic University of Norway  
Tromso, Norway

**For correspondence:** *Grjibovski Andrej*; e-mail: Andrej.Grijibovski@gmail.com

**Conflict of interest.** The authors declare no conflict of interest.

**Funding.** The study had no sponsorship.

**Received:** 13.06.2017.

In this article, we have presented basic principles of the parametric criteria using for comparing two and three paired groups. Practical examples are presented with step-by-step algorithm on how to use SPSS and STATA software for paired t-test and one-way repeated measures analysis of variance (RM-ANOVA) with subsequent interpretation of the outputs.

**Keywords:** statistical analysis, paired samples, paired t-test, one-way repeated measures analysis of variance, SPSS, STATA