

УДК 311.2:61

КРАТКИЕ РЕКОМЕНДАЦИИ ПО ОПИСАНИЮ, СТАТИСТИЧЕСКОМУ АНАЛИЗУ И ПРЕДСТАВЛЕНИЮ ДАНЫХ В НАУЧНЫХ ПУБЛИКАЦИЯХ

© 2011 г. ^{1,2}Т. Н. Унгуряну, ²⁻⁴А. М. Гржибовский¹ Управление Роспотребнадзора по Архангельской области, г. Архангельск² Архангельская международная школа общественного здоровья,
Северный государственный медицинский университет, г. Архангельск³ Норвежский институт общественного здравоохранения, г. Осло, Норвегия⁴ Университет г. Тромсё, Норвегия

Цель настоящей работы — предоставить авторам алгоритм выбора способа описания и анализа данных в зависимости от их типа и распределения. Алгоритм включает лишь наиболее часто встречающиеся в рукописях, представляемых в журнал, методы статистической обработки данных.

Прежде чем описывать полученные в ходе исследования данные, необходимо определить, количественными или качественными они являются. Количественные признаки имеют реальное числовое выражение, например, масса и длина тела, длительность госпитализации, количество выкуриваемых сигарет в день и т. д. Качественные признаки характеризуют свойства объекта, например, пол, семейное положение, уровень образования, стадия болезни и т. д. [1, 7]. С качественными переменными, даже если они закодированы исследователем с помощью цифр, невозможно производить арифметические действия.

Выбор мер для описания количественных данных зависит от типа их распределения. Поэтому, перед тем как описывать количественные данные, нужно проверить их распределение. Проверка распределения на основании собранных данных необходима, когда исследователю неизвестно распределение количественного признака в генеральной совокупности. В этом случае тип распределения можно определить с помощью критериев Шапиро — Уилка для выборок с числом наблюдений 50 и менее и Колмогорова — Смирнова для больших выборок. При использовании данных критериев за нулевую принимается гипотеза о том, что изучаемое распределение не отличается от нормального. Если достигнутый уровень значимости (p) при проверке гипотезы будет меньше, чем критический (чаще всего принимаемый за 0,05), то нулевая гипотеза отклоняется, а это значит, что распределение отличается от нормального. Если достигнутый уровень значимости будет больше, чем 0,05, то распределение является нормальным. Если анализируются две группы и более, тип распределения данных определяется для каждой группы. Целесообразно проверять распределение не только с помощью вышеперечисленных статистических критериев, но и графически — с помощью квантильных диаграмм, так как при небольших выборках статистические критерии могут не выявить даже выраженных отклонений от нормальности, тогда как при наличии больших (тысячи наблюдений) выборок даже незначительное отклонение от нормальности будет статистически значимым.

Для описания количественных данных, имеющих нормальное распределение, используются среднее арифметическое (M) и стан-

Авторы дают краткие рекомендации по описанию данных, их анализу и представлению в научных публикациях (особое внимание уделяется методам определения типа распределения данных и выбору соответствующих методов статистического анализа в зависимости от типа распределения и количества сравниваемых групп); использованию следующих статистических критериев и методов: тестов Шапиро — Уилка и Колмогорова — Смирнова, одновыборочного и двухвыборочного критериев Стьюдента, критериев Манна — Уитни, Вилкоксона, Краскела — Уоллиса, Фридмана, хи-квадрат Пирсона, Кокрена, Мак-Нимара, точного критерия Фишера, однофакторного дисперсионного анализа и дисперсионного анализа повторных измерений, а также ссылки на литературу, где вышеперечисленные тесты описаны более детально. Авторы надеются, что эти рекомендации помогут исследователям правильно использовать статистические методы и улучшить представление результатов в своих работах.

Ключевые слова: описание данных, статистический анализ, представление данных.

дартное отклонение (SD), которые целесообразно представлять в виде $M(SD)$. Использование стандартной ошибки среднего арифметического (m) в качестве меры рассеивания не рекомендуется, так как она характеризует не разброс данных вокруг среднего значения [2], а лишь точность измерения среднего. Как правило, исследования проводятся на выборочных данных, а результаты экстраполируются на всю популяцию или популяционную группу, поэтому наряду с выборочными оценками желательно представлять 95 % доверительный интервал для популяционной средней. Доверительный интервал показывает, в каких пределах с 95 % доверительной вероятностью находится истинное среднее популяции [4]. Если количественные данные не подчиняются закону нормального распределения, для их описания применяются медиана (Me) и процентиля. Чаще рассчитываются 25-й (P_{25}) и 75-й (P_{75}) процентиля, которые также

называются квантилями (Q_1 и Q_3 соответственно). Для интервальной оценки медианы также можно рассчитать доверительный интервал, по которому можно судить, в каких пределах находится медианное значение для популяции при заданном уровне доверительной вероятности (чаще всего 95 %).

Для описания качественных данных используются частоты и доли (в %), с которыми те или иные значения качественных признаков встречаются в выборке. Как и для выборочной средней, для выборочной доли (пропорции) тоже необходимо рассчитывать 95 % доверительный интервал. По доверительному интервалу для пропорции можно судить о том, в каких пределах находится значение изучаемой доли в генеральной совокупности. Существует много способов расчета доверительных интервалов для долей, которые детально описаны в [5].

Выбор статистического критерия для оценки различий между средними значениями в группах зависит

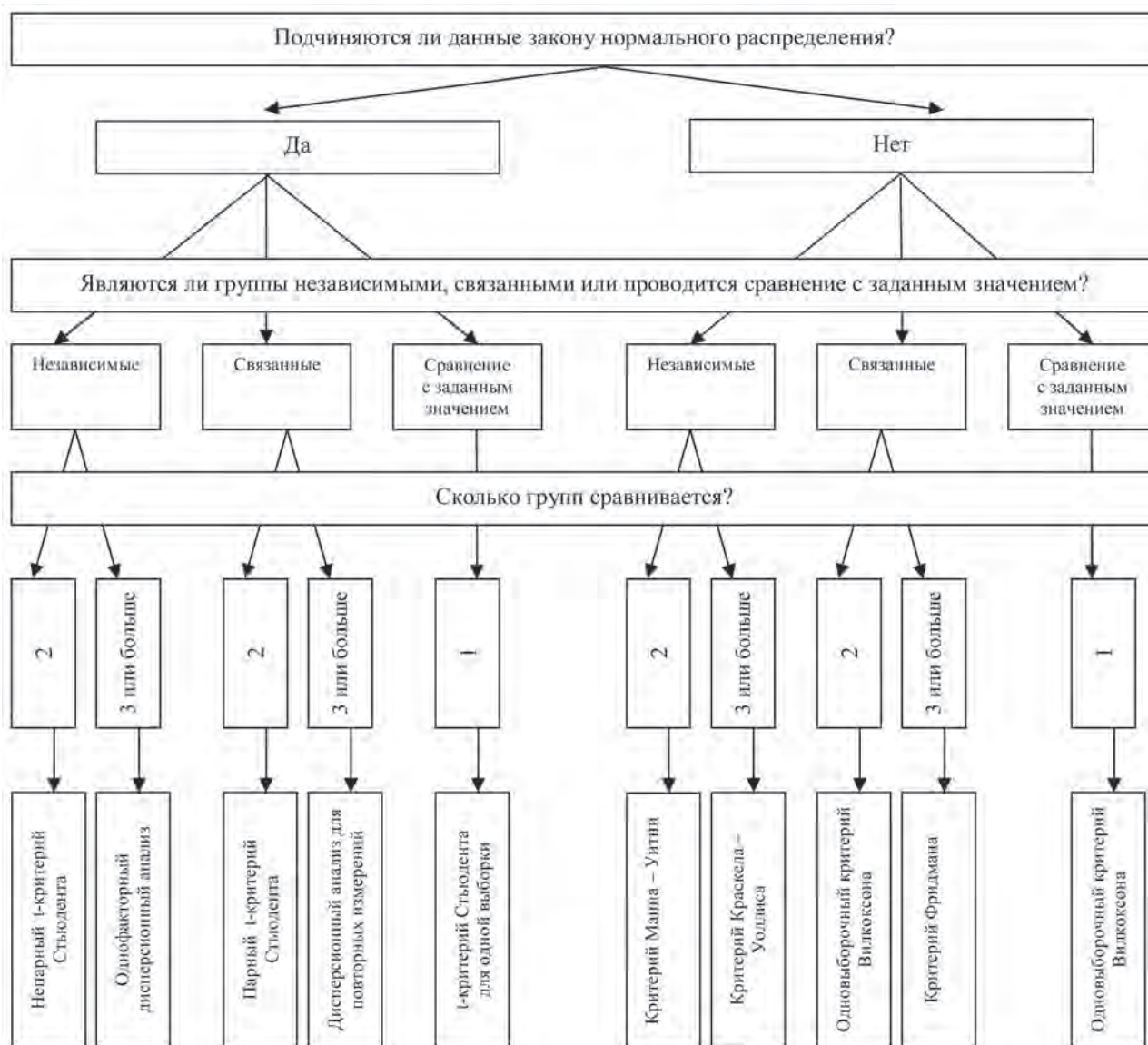


Рис. 1. Алгоритм выбора статистического критерия для анализа количественных данных [3]

от количества сравниваемых групп и типа распределения данных (рис. 1). При нормальном распределении данных используются параметрические критерии, в то время как при асимметричном распределении — непараметрические.

Для сравнения средних значений в двух независимых группах (например, в группе мужчин и группе женщин), в случае, если данные в обеих группах подчиняются закону нормального распределения, используется двухвыборочный (непарный) критерий Стьюдента [2, 7]. Если распределение данных в обеих группах или хотя бы в одной группе является асимметричным, то для сравнения средних значений (в данном случае медианных) следует использовать непараметрические критерии — Манна — Уитни или двухвыборочный Вилкоксона.

Для сравнения средних значений в двух парных выборках (например, обследование проводилось до воздействия и после него) при нормальном распределении данных в обеих группах применяется одновыборочный (парный) критерий Стьюдента. Если распределение данных хотя бы в одной группе является асимметричным, то сравнение средних значений двух парных групп проводят с помощью одновыборочного критерия Вилкоксона [2, 7].

Если необходимо сравнить средние значения количественного признака в трех и более независимых группах (например, уровень артериального давления в трех группах пациентов с разными методами лечения), то при нормальном распределении данных во всех группах используется однофакторный дисперсионный анализ (ANOVA). В случае, если нужно сравнить средние значения в трех и более зависимых выборках (например, уровень артериального давления в группе пациентов до воздействия, через 10 минут и через 30 минут после воздействия), при нормальном распределении количественного признака во всех группах используется однофакторный дисперсионный анализ для повторных измерений (Repeated Measures ANOVA).

При интерпретации результатов дисперсионного анализа следует принимать во внимание, что выявление статистически значимых различий говорит только о том, что различия между средними существуют, но не о том, какие из групп различаются между собой. Поэтому следующим шагом после того, как по данным дисперсионного анализа выявлены статистически значимые различия, являются апостериорные сравнения. Апостериорные сравнения представляют собой попарные сравнения изучаемых групп для обнаружения различий между ними. Существует более 18 критериев для апостериорных сравнений, но наиболее простым и популярным способом попарных сравнений является поправка Бонферрони, которая рекомендуется при небольшом (до 4–5) количестве сравнений.

Для сравнения трех и более независимых групп, в которых данные не подчиняются закону нормального распределения, следует применять критерий Краскела — Уоллиса, являющийся непараметрическим аналогом однофакторного дисперсионного анализа. При сравнении трех и более связанных (зависимых) выборок, данные в которых не подчиняются закону нормального распределения, необходимо использовать критерий Фридмана [2]. Эти непараметрические критерии рекомендуется использовать даже в том случае, если хотя бы в одной из сравниваемых групп распределение данных является асимметричным.

Как и дисперсионный анализ, критерии Краскела — Уоллиса и Фридмана позволяют в целом оценить наличие различий между группами и не позволяют судить о том, какие группы различаются между собой. Поэтому после обнаружения статистически значимых различий между группами в целом, необходимо провести попарные сравнения. Для попарного сравнения независимых выборок можно использовать критерий Манна — Уитни или двухвыборочный критерий Вилкоксона, а для попарного сравнения зависимых выборок — одновыборочный критерий Вилкоксона. Как было сказано выше, данные критерии предназначены для сравнения только двух групп. Поэтому при их использовании для попарных сравнений необходимо рассчитать новый критический уровень статистической значимости для оценки справедливости нулевой гипотезы. Для расчета нового критического уровня статистической значимости нужно 0,05 разделить на число попарных сравнений. Например, если сравниваются попарно три группы между собой (1 и 2; 2 и 3; 1 и 3), то новый критический уровень статистической значимости составит: $0,05 \div 3 = 0,017$, а если попарно сравниваются четыре группы, то новый критический уровень статистической значимости составит: $0,05 \div 6 = 0,0085$.

Если задачей является сравнение качественных, номинальных переменных (долей), то, прежде чем выбрать статистический критерий, необходимо определиться с количеством сравниваемых групп (рис. 2). При сравнении частоты встречаемости признака или заболевания в одной выборке с заданным значением можно воспользоваться критерием Z [3].

Сравнение двух независимых групп номинальных данных (например, сравнение удельного веса курящих среди мужчин и женщин) проводится с использованием критерия хи-квадрат (χ^2) Пирсона [2, 7]. Если ожидаемое число наблюдений в любой из ячеек четырехпольной таблицы 2×2 окажется меньше 5, то следует применять точный критерий Фишера. Для сравнения долей в двух зависимых выборках используется критерий Мак-Нимара.

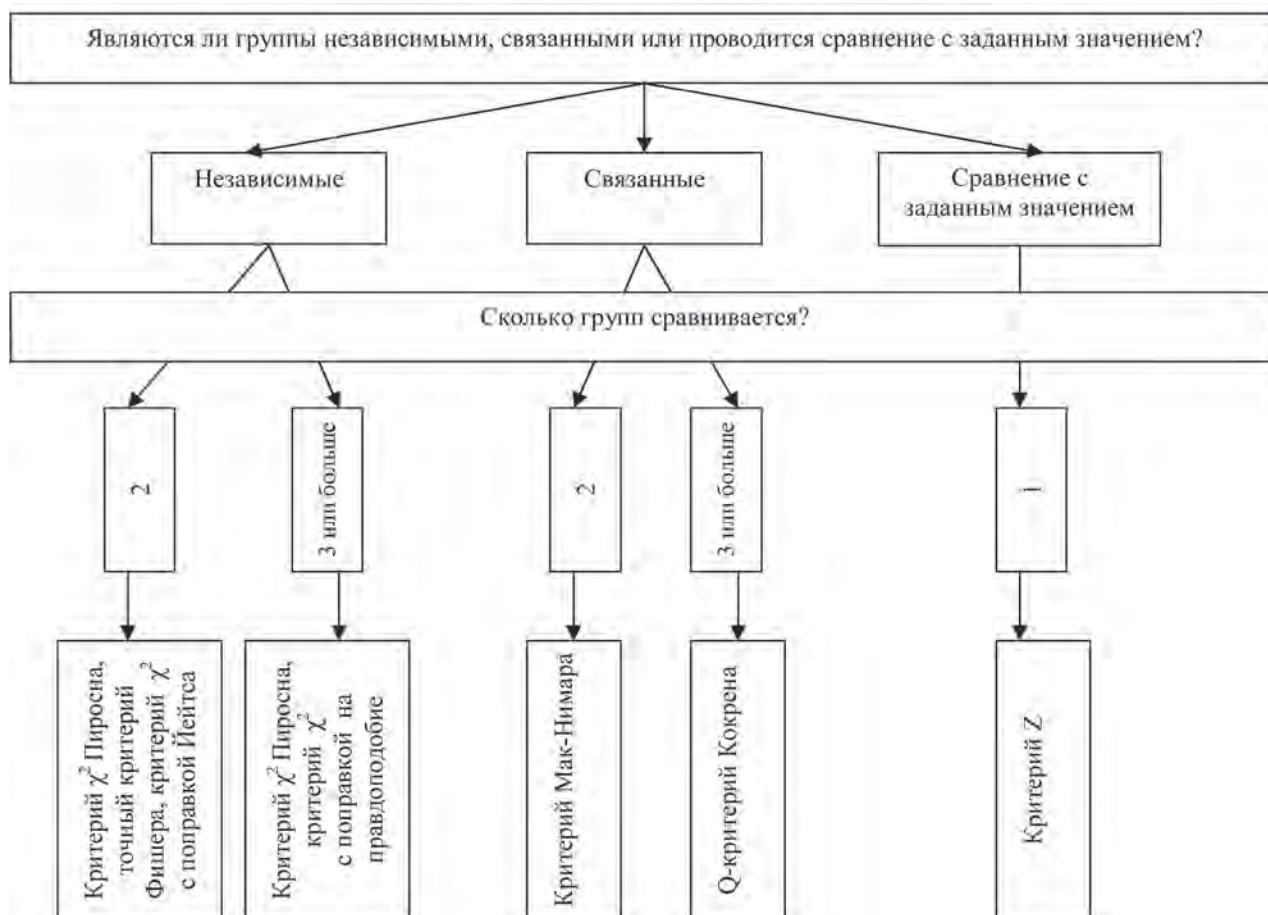


Рис. 2. Алгоритм выбора статистического критерия для анализа качественных данных [3]

Если нужно сравнить качественные данные в трех и более независимых группах, то также можно воспользоваться критерием χ^2 Пирсона. При этом, если в многопольной таблице, например, 2×3 , 3×3 и т. д., доля ячеек с ожидаемым числом наблюдений будет более 20 % или хотя бы одно ожидаемое значение окажется менее единицы, то для расчета критерия χ^2 рекомендуется объединить группы, чтобы увеличить количество наблюдений в клетках таблицы. При сравнении качественных данных в трех и более зависимых группах используется Q-критерий Кокрена.

При представлении результатов статистического анализа данных рекомендуется указывать точные значения критерия и точное значение достигнутого уровня статистической значимости, округлив его до трех знаков после запятой, например $p = 0,032$. Не рекомендуется указывать достигнутые уровни статистической значимости как $p < 0,05$, $p < 0,01$ или $p > 0,05$. Исключение составляют очень малые значения p , которые можно представлять, как $p < 0,001$. Для параметрических и некоторых непараметрических критериев также рекомендуется приводить число степеней свободы (df , degrees of freedom) (таблица). Следует всегда помнить, что выявление статистически значимых различий еще не говорит о

том, что эти различия истинны. Мы не рекомендуем использовать термин «достоверные различия», так как наличие статистически значимых результатов не означает их достоверности.

При анализе количественных данных, кроме сравнения средних значений в группах, иногда стоит задача выявления взаимосвязи между переменными. Если обе переменные (x и y) количественные, подчиняются закону нормального распределения и между ними существует линейная зависимость, то для выявления взаимосвязи между этими переменными используется коэффициент линейной корреляции Пирсона. Но если распределение переменных является асимметричным или связь между переменными нелинейная, то корреляционный анализ выполняется с помощью коэффициента ранговой корреляции Спирмена. Коэффициент Спирмена также можно использовать для выявления связи, когда одна из переменных (x или y) является порядковой (ранговой) [7]. Предъявляя результаты корреляционного анализа, обязательно нужно представлять не только значение коэффициента корреляции (с указанием, что это за коэффициент), но и статистическую значимость выявленной связи между переменными. Исследователи также

Рекомендации по представлению результатов статистических критериев

Критерий	Форма представления	Пример
Непарный Стьюдента	t, df, p	$t = -3,45, df = 234, p = 0,002$
Парный Стьюдента	t, df, p	$t = 1,86, df = 128, p = 0,068$
Манна – Уитни	U, Z, p	$U = 7843, Z = -1,23, p = 0,032$
Двухвыборочный Вилкоксона	W, Z, p	$W = 2543, Z = -2,64, p = 0,004$
Одновыборочный Вилкоксона	W, Z, p	$W = 1276, Z = -3,32, p = 0,006$
Дисперсионный анализ	F, df, p	$F_{2,105} = 2,87, p = 0,088$
Краскела – Уоллиса	H, df, p	$H = 8,65, df = 3, p = 0,123$
Фридмана	Fr, df, p	$Fr = 0,21, df = 2, p = 0,974$
Z	Z, p	$Z = 13,2, p = 0,034$
χ^2 Пирсона	χ^2, df, p	$\chi^2 = 1,85, df = 1, p = 0,543$
Мак-Нимара	χ^2, df, p	$\chi^2 = 5,61, df = 1, p = 0,003$
Q Кокрена	Q, df, p	$Q = 2,37, df = 2, p = 0,305$
Коэффициент корреляции Пирсона	r_p, p	$r_p = 0,75, p = 0,001$
Коэффициент корреляции Спирмена	r_s, p	$r_s = -0,82, p = 0,008$

Примечание. df (degrees of freedom) – число степеней свободы, p – достигнутый уровень статистической значимости.

могут рассчитать доверительные интервалы для коэффициентов корреляции.

Корреляционный анализ позволяет определить силу и направление связи между переменными. Если стоит задача спрогнозировать (предсказать), как изменятся значения зависимой переменной при изменении независимой переменной, используется регрессионный анализ. Простой линейный регрессионный анализ применяется, если обе переменные являются количественными и между ними существует линейная зависимость. Если имеется несколько независимых переменных, с помощью которых прогнозируется зависимая переменная, то применяется многомерный линейный регрессионный анализ, представляя результаты которого, необходимо указывать нескорректированные и скорректированные значения коэффициентов и их доверительные интервалы, а также значение коэффициента детерминации регрессионной модели. Логистический регрессионный анализ проводят, если исход является бинарной величиной (например, наличие/отсутствие болезни) [7]. Представляя результаты логистического регрессионного анализа, необходимо указывать отношение шансов (ОШ) и 95 % доверительные интервалы для ОШ. Указывая доверительные интервалы, приводить уровень значимости (p) необязательно.

Более редкие методы анализа, такие, например, как факторный анализ, кластерный, дискриминантный, анализ нелинейных зависимостей, многомерный дисперсионный анализ и т. п., не рассматриваются

по причине их редкой встречаемости в рукописях, представляемых в журнал.

Авторы также рекомендуют исследователям ознакомиться с примерами наиболее часто встречающихся ошибок представления данных в научных публикациях [3], статьями рубрики «Практикум» журнала «Экология человека» за 2008 год (№ 1–12), где подробно рассмотрены наиболее часто встречающиеся в медицинской литературе статистические критерии, правила выбора наиболее подходящего способа обработки данных в зависимости от имеющегося материала, алгоритмы применения выбранных статистических процедур с использованием пакета программ SPSS и интерпретации полученных результатов, а также полным текстом Единых требований к рукописям, представляемым в научно-практический журнал «Экология человека» [6].

Список литературы

1. Банержи А. Медицинская статистика понятным языком: вводный курс / А. Банержи. – М.: Практическая медицина, 2007. – 287 с.
2. Гланц С. Медико-биологическая статистика / С. Гланц. – М.: Практика, 1998. – 459 с.
3. Гржибовский А. М. Выбор статистического критерия для проверки гипотез / А. М. Гржибовский // Экология человека. – 2008. – № 11. – С. 48–57.
4. Гржибовский А. М. Использование статистики в Российской биомедицинской литературе / А. М. Гржибовский // Экология человека. – 2008. – № 12. – С. 55–64.
5. Гржибовский А. М. Однофакторный линейный регрессионный анализ / А. М. Гржибовский // Экология человека. – 2008. – № 10. – С. 55–64.

6. *Единые* требования к рукописям, представляемым в научно-практический журнал «Экология человека» // Экология человека. — 2010. — № 12. — С. 55–60.

7. *Петри А.* Наглядная медицинская статистика / А. Петри, К. Сэбин. — М. : ГЭОТАР-Медиа, 2009. — 168 с.

BRIEF RECOMMENDATIONS ON DESCRIPTION, ANALYSIS AND PRESENTATION OF DATA IN SCIENTIFIC PAPERS

^{1,2}T. N. Unguryanu, ²⁻⁴A. M. Grjibovski

¹*Administration of Federal Service for Surveillance on Consumer Rights Protection and Human Wellbeing in the Arkhangelsk region, Arkhangelsk, Russia*

²*International School of Public Health, Northern State Medical University, Arkhangelsk, Russia*

³*Norwegian Institute of Public Health, Oslo, Norway*

⁴*University of Tromsø, Norway*

The authors present brief recommendations on data description, data analysis and data presentation in scientific papers with special emphasis on checking the data for normality

and selection of appropriate methods of statistical analysis depending on distribution of data and number of groups. Brief guidelines on the use of Shapiro-Wilk and Kolmogorov-Smirnov tests as well as one-sample and two-samples t-tests, one-way and repeated measures ANOVA, Mann-Whitney, Wilcoxon, Kruskal-Wallis, Friedman, Pearson's chi-squared, Cochran, MacNemar and Fisher exact tests are presented. References to more detailed description of each of the above-mentioned tests are also given. The authors hope that these brief guidelines will assist the authors and help to improve the use and presentation of the results of statistical procedures in the manuscripts submitted to the journal.

Key words: description, data analysis, data presentation.

Контактная информация:

Унгуриану Татьяна Николаевна — кандидат медицинских наук, магистр общественного здоровья, главный специалист-эксперт Управления Роспотребнадзора по Архангельской области, доцент Северного государственного медицинского университета

Адрес: 163061, г. Архангельск, ул. Гайдара, д. 24

Тел. (8182) 21-04-61

E-mail: unguryanu_tn@mail.ru