

В.П. Леонов

ТРЕБОВАНИЯ И РЕКОМЕНДАЦИИ ПО ОПИСАНИЮ И ИСПОЛЬЗОВАНИЮ СТАТИСТИЧЕСКОГО АНАЛИЗА В ИССЛЕДОВАНИЯХ

✉ В.П. Леонов, к.т.н., доцент, НЦ БИОСТАТИСТИКА, редактор сайта БИОМЕТРИКА (<http://www.biometrica.tomsk.ru/>),
e-mail: leo.biostat@mail.com

V.P. Leonov

REQUIREMENTS AND RECOMMENDATIONS FOR DESCRIPTION AND USAGE OF STATISTICAL ANALYSIS IN STUDIES

✉ V.P. Leonov, CTS, assistant professor, NC Biostatistics, editor of the site 'Biometrica' (<http://www.biometrica.tomsk.ru/>),
e-mail: leo.biostat@mail.com

Приводимые ниже редакционные требования и рекомендации по описанию и использованию статистических аспектов, обусловлены как целями авторских публикаций, так и задачами самого журнала. В частности, основной целью высокого уровня описания статистического анализа в статье является генерация у читателей необходимого доверия к публикуемым результатам исследований. Реализация авторами статей требований в описании результатов своих научных исследований, проводимых с учётом наших требований и рекомендаций, должна приводить к появлению у читателей трёх основных последствий:

1. Ясное понимание и доверие к результатам исследования;
2. Рост вероятности реализации в своей практике описанных результатов исследования;
3. Продолжение подобного исследования с использованием описанных методов и результатов, принося к собственной базе данных.

Соответственно высокий уровень описания этих статистических методов авторами статей должен приводить к повышению привлекательности чтения данного журнала и увеличению количества направляемых для их публикации рукописей статей с результатами продуктивных исследований. При этом авторы, присылая в редакцию журнала рукопись статьи, принимают на себя следующее обязательство: «В случае сомнения у рецензентов журнала в корректности описываемых результатов статистического анализа, мы согласны предоставлять рецензенту электронный вариант базы данных в формате пакета EXCEL для проверки корректности этих результатов». Обращаем внимание на то, что данное продуктивное требование было представлено в ведущих журналах США и Англии уже в конце 19 века. Отмечаем, что рецензент, продуктивный по статистическому анализу, при проверке корректности результатов по высланной авторами базе данных, также сообщит и рекомендации

по использованию в данном исследовании и дополнительных методов статистического анализа.

Для оперативного освоения специфики описания деталей статистического анализа, проводимого авторами статей, им необходимо детально познакомиться как с изданиями по описанию методов и результатов статистического анализа, так и с примерами типичных ошибок описания этих результатов в статьях и диссертациях. Информация по двум этим направлениям приведена ниже.

Учитывая весьма большой набор использования самых разных методов статистического анализа, и большое количество статистических гипотез и критериев, рекомендуем авторам рукописей статей прочитать основную по данному направлению книгу «Как описывать статистику в медицине. Руководство для авторов, редакторов и рецензентов». Авторы: Томас А. Ланг, Мишель Сесик. Перевод с английского В.П. Леонова. Изд-во «Практическая медицина», Москва, 2016. 477 с. (см. URL: http://www.biometrica.tomsk.ru/Lang_Secic_2015.pdf). Отметим, что объёмы подобных описаний являются весьма большими ввиду наличия многих методов статистического анализа и критериев, параметров исследования, характеристик описания выбора методов анализа, таблиц, графиков и т.д. Поэтому не случайно объём упомянутой книги «Как описывать статистику в медицине. Руководство для авторов, редакторов и рецензентов» содержит 477 страниц. Это издание содержит 5 частей, в которых приведена 21 глава, а также 5 Приложений, Библиография и подробный Предметный указатель. Отметим, что при переводе данного издания, издательству в дополнение переводчиком был предложен и большой список «Библиографии книг по статистическому анализу» на русском языке. Издательство согласилось опубликовать русскоязычную Библиографию. Однако, как оказалось, организатор печати этого издания по какой-то причине не ввёл в печать этот дополнительный вари-

ант Библиографии. Аналогично при переводе книги «Наглядная медицинская статистика: учебное пособие». Авторы: А. Петри, К. Сэбин. Перевод с английского В.П. Леонова. – 3-е издание, переработанное и дополненное. М.: ГЭОТАР-Медиа, 2015. 216 с. (См. URL: http://www.biometrica.tomsk.ru/Medstat_small.pdf) было предложено подобное «Приложение к русскому изданию: Библиография от научного редактора», содержащее 459 изданий из 433 русскоязычных книг и статей. При печати этого издания данное приложение было опубликовано. Также предлагаем использовать и книгу «Прикладная математическая статистика. Для инженеров и научных работников». М.: ФИЗМАТЛИТ, 2006. 816 с., в которой приведено 5 больших глав, а также 638 изданий в «Списке литературы». Эти три весьма полезные издания рекомендуем авторам рукописей статей использовать в своих исследованиях. Их можно приобрести в книжных магазинах. Для чего сделать поиск этих книг в Интернете либо найти их для чтения в доступных библиотеках.

Ввиду отсутствия у авторов подобных изданий, во многих публикуемых статьях и диссертациях по медицине и биологии обнаруживается очень немало ошибок использования и описания статистических аспектов проводимого исследования. Поэтому у читателей таких публикаций возникают сомнения в истинности приводимых авторами результатов исследования. Примеры весьма критических ошибок описания статистического анализа приведены в статьях «Применение методов статистики в кардиологии (по материалам журнала «Кардиология» за 1993-1995 гг. Журнал «Кардиология, том 38, 1998, № 1. (См. URL: http://www.biometrica.tomsk.ru/Cardiology_1998_1.pdf), «Статистика в кардиологии. 15 лет спустя» (См. URL: http://www.biometrica.tomsk.ru/stat_cardiol.htm), «Доказательная или сомнительная? Медицинская наука Кузбасса: статистические аспекты» (См. URL: <http://www.biometrica.tomsk.ru/kuzbass1.htm>), и в разделе «КУНСТКАМЕРА» сайта БИОМЕТРИКА по адресу <http://www.biometrica.tomsk.ru/kk.htm> немало примеров ошибок описания использования статистических методов в публикациях по медицине, приведено в докладе «Почему и как надо учить медиков статистике?» (См. URL: http://www.biometrica.tomsk.ru/erevan_4.html) на международной конференции по доказательной медицине в Ереване 18-20.10.2012 г. Причём в этом докладе также приведены названия 208 книг и статей по методам статистического анализа. Аналогичные примеры ошибок отмечены и в докладе «Цели, возможности, и проблемы использования биостатистики в доказательной медицине» (См. URL: http://www.biometrica.tomsk.ru/Leonov_Erevan_2015.pdf) на международной Конференции в Ереване «От доказательной медицины к доказательному здравоохранению» (24 - 26 сентября 2015 года).

Детализация описания результатов статистического анализа представляется как в конкретных числовых значениях, так и в конкретных их описаниях. Например, в наличии значений средней величины количественного признака, доверительного интервала, стандартного отклонения среднего значения, дисперсии, медианы, моды, нижнего и верхнего квартилей, процентилей, конкретных значений статистического критерия, достигнутого уровня статистической значимости, критического уровня статистической значимости Р, отношения шансов, коэффициента сопряжённости, отношения рисков, процентных отношений, априорных вероятностей, байесовского коэффициента, чувствительности, специфичности, коэффициентов корреляции, коэффициентов автокорреляции, квадрата корреляционного отношения, размерные и безразмерные регрессионные коэффициенты, собственные значения многомерных методов, параметрических и непараметрических критериев и т.п. При этом приводить как правильные названия использованных методов статистического анализа и терминов, так и правильные названия используемых статистических символов. Следует приводить и названия использованных статистических пакетов с указанием номера пакета. Данная детализация описания этих результатов вызовет большое доверие к результатам исследования у тех читателей, которые владеют аналогичным уровнем знания по статистическому анализу. А у читателей, не владеющих таким уровнем знаний, это будет способствовать повышению знаний по статистическому анализу и проведению подобных исследований, либо самостоятельному продолжению этого исследования.

ОПИСАНИЕ ИСПОЛЬЗОВАННЫХ МЕТОДОВ И РЕЗУЛЬТАТОВ СТАТИСТИЧЕСКОГО АНАЛИЗА

1. В начале раздела «Материал и методы» следует подробно описать технологию сбора анализируемой базы данных, так и объём анализируемой базы данных. В частности, описать количество наблюдений, число количественных и число качественных (группирующих) признаков. При этом сообщить наличие или отсутствие доли пропущенных значений анализируемых признаков. Например, если база данных содержала 80 наблюдений и 10 признаков, и при этом были все $80 \times 10 = 800$ значений, то количество пропусков в базе данных равно нулю. Если же в такой базе данных не были введены 80 значений, то указывается доля пропущенных значений, равная $80/800 = 10\%$ и т.п. Если исследование было рандомизированным, сообщить детали рандомизации.

2. Привести названия статистических пакетов с указанием номеров версий этих пакетов. Обращаем внимание читателей данных требований на то, что

содержание получаемых результатов различных видов статистического анализа в принципе зависит от используемого статистического пакета. Т.е. в разных статистических пакетах, в том числе в пакетах с одним и тем же названием, но с различающимися вариантами (англоязычные или русскоязычные, а также с разными номерами пакетов), получаемые результаты одинаковых статистических методов могут содержать отличающиеся форматы и описания этих результатов. Поэтому авторам публикуемых результатов статистического анализа следует понимать, что приводимые далее примеры описания различных методов статистического анализа не являются неизменными.

3. Описать цели использования применяемых методов статистического анализа. Привести правильные названия использованных методов статистического анализа. При этом уточняйте, какие методы статистического анализа и какие статистические критерии использовались для каждой конкретной цели исследования. Для упоминаемых статистических методов важно приводить соответствующие литературные ссылки, причём желательно на литературу на русском языке. Если целью исследования была проверка конкретных гипотез, их следует ясно описать. Например, о равенстве соответствующих групповых средних либо о равенстве коэффициентов корреляции и т.п.

4. Для непрерывных количественных признаков следует обязательно произвести проверку нормального распределения, при этом используя не один лишь критерий, например, критерий Колмогорова-Смирнова, но также и ещё, как минимум, один критерий. Например, критерии Шапиро-Уилка, Крамера – фон Мизеса, Андерсона-Дарлинга, Шапиро-Франча, Лиллиефорса, Локка-Спурье, Д’Агостино, Локка-Спурье и т.д. Причём для небольших объёмов, порядка 20 – 100 наблюдений, желательно использовать мощный критерий Шапиро-Уилка. Если авторы для разных целей, например, для нормализации количественного признака применяли преобразование признака исходных данных, то следует точно описать процедуры этого преобразования. Результаты проверки нормальности распределения должны содержать как название использованного статистического критерия, так и значение достигнутого уровня статистической значимости. Например: «Для проверки нормальности распределения всех описанных количественных признаков использовали критерий Колмогорова-Смирнова и Шапиро-Уилка. Все количественные признаки в сравниваемых группах не имеют нормального распределения».

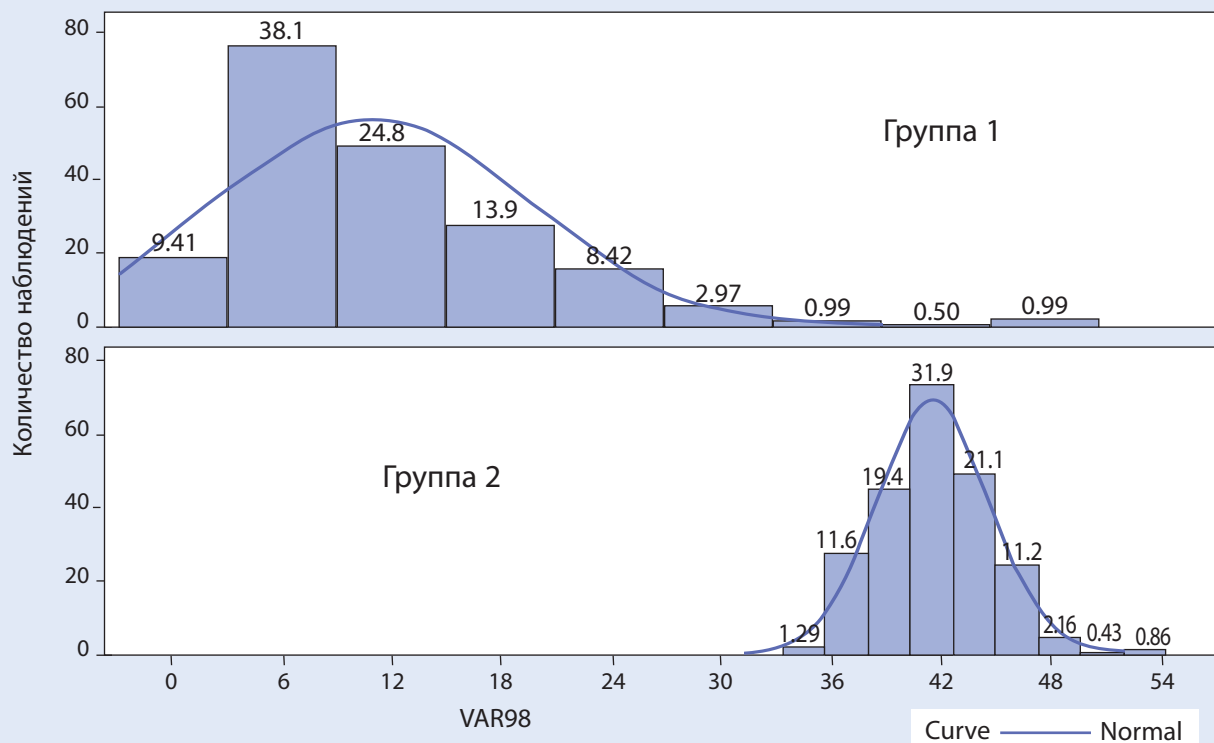
5. Для обсуждаемых количественных признаков, имеющих нормальное распределение, приводятся количество анализируемых значений (N), величина среднего значения (M) и величина стандартного отклонения (SD) в формате $M \pm SD$. Если конкрет-

ный количественный признак не имеет нормального распределения, то его параметры приводятся в виде значения медианы и двух квартилей: $Me [Q1; Q3]$. Данные значения приводятся с точностью до второго знака после запятой. При этом по результатам проверки всех статистических гипотез необходимо указывать метод статистического анализа, в котором получен данный уровень статистической значимости. И приводить этот уровень с тремя цифрами после запятой (например, $p=0,018$), а не в виде неравенств типа $p>0,05$ или $p<0,05$). Если же подобные ненулевые значения располагаются после запятой и после 3-4-х и более нулей, то следует эти значения приводить в формате типа $p = 4,83 \cdot 10^{-4}$. Эти значения следует приводить в таблицах, содержащих данные величины для всех анализируемых количественных переменных. Также при необходимости авторы могут представлять подобные значения и в виде графиков. Если же в анализе используются доверительные интервалы, то следует обязательно указывать тип доверительного интервала (двусторонний или односторонний), и значение доверительной вероятности.

6. Учитывать соответствие приводимых утверждений на основе результатов статистического анализа с величиной уровня статистической значимости проверенных статистических гипотез. В частности, при обсуждении результатов статистического анализа как первоначального варианта вскрытых зависимостей, можно использовать критический уровень статистической значимости в 5%, т.е. приводить выражение «Критический уровень статистической значимости $p=0,05$ ». Если же получаемый результат статистического анализа приводится как заключение об использовании в реальной биомедицинской практике, то желательно использовать критический уровень статистической значимости в 1%, или даже в 0,5%. т.е. приводить выражения « $p = 0,01$ » или « $p = 0,005$ ». При этом нельзя использовать выражения типа «Уровень достоверности $p = 0,05$ ». Поскольку выражения со словом «достоверно» являются принципиально ошибочными, не соответствующим «ГОСТ Р 50779.10-2000 «Статистические методы. Вероятность и основы статистики. Термины и определения».

7. В проведении статистического анализа в исследовании чаще всего используются те методы анализа, в которых совместно анализируются 2 или более признаков. Например, при анализе связи пары качественных признаков используют анализ таблицы сопряжённости. При анализе двух количественных признаков используют методы корреляционного или регрессионного анализа. А при анализе одной группирующей переменной и одной количественной переменной применяются разные методы по сравнению групповых параметров. В частности, сравниваются групповые средние, групповые дисперсии, груп-

Гистограммы VAR98 в группах 1 и 2 признака GRUPPA



повые медианы и т.д. Похожие анализы производятся и при использовании более двух признаков. Обычно авторы публикаций считают единственной основной целью использования подобных методов лишь получение ответа на гипотезы о статистически значимой взаимосвязи этих двух и более признаков. Однако при наличии статистически значимой взаимосвязи этих признаков, необходимо также и описание структуры каждой из такой связи. В частности, какой конкретно вид имеет установленная связь. Кроме того, необходимо устанавливать какие комбинации подгрупп или признаки доминируют в этих парных или многомерных связях. Естественно, что для реализации этой редакционной рекомендации, авторам рукописи статьи необходимо владеть достаточным уровнем теории этих статистических методов. Если же такого уровня знания нет, то следует обращаться за помощью к специалистам по статистике в медицине и биологии. Ниже, в отдельных группах рекомендаций, приведены фрагменты таких желательных описаний структуры парных или многомерных взаимосвязей.

8. По группирующим признакам, которые используются при сравнениях групповых параметров, обязательно описать количество группировок и названия этих группировок. При сравнении группирующих параметров количественных признаков сообщить аргументацию выбора конкретных статистических критериев. Например, «Ввиду отсутствия нормального распределения и неравенства групповых

дисперсий, для сравнения трёх групповых средних использовался критерий Уилкоксона».

9. Для количественного признака в сравниваемых группах рекомендуется привести графическое отображение гистограмм распределения в этих группах с наличием кривой нормального распределения, что может отразить как вид самого распределения, так и характер взаимного расположения сравниваемых групп. Ниже приведён пример такого графика, на котором видны типы распределения. Также кривая линия отражает теоретическое нормальное распределение по параметрам признака данной группы. Видно и взаимное расположение значений количественного признака в двух сравниваемых группах. Решение о публикации подобных графиков в рукописи статьи принимают сами авторы публикации.

10. При сравнении группирующих средних с помощью параметрического t-критерия Стьюдента или классического дисперсионного анализа ANOVA, следует обязательно производить проверку двух условий корректного использования данных методов: нормальности распределения во всех сравниваемых группах и равенство генеральных (популяционных) дисперсий по всем группам. При этом по результатам проверки статистических гипотез необходимо указывать метод статистического анализа, в котором получен данный уровень статистической значимости. А приводить этот уровень с 3-4 цифрами после запятой (например, $p=0,075$ или $p=0,0012$), а не в виде нера-

венств типа $p > 0,05$ или $p < 0,05$. Если же одно или оба этих условий дают отрицательный результат, то следует использовать соответствующие непараметрические критерии. Например, критерии Ван дер Вардена, Вилкоксона, Краскела-Уоллиса, Вальда-Вольфовица, Манна-Уитни, ранговый критерий Фридмана, Левиса, медианный критерий, Q-критерий Кокрена, H-критерий и т.д.

Ниже приводим один из примеров такого результата по сравнению групповых параметров признака SR114 с использованием дисперсионного анализа (ANOVA) и критериев Kruskal-Wallis, медианный критерий, Ван-дер-Вардена, Siegel-Tukey (Сиджела-Тьюки: критерий сравнения дисперсий). Результат ANOVA при сравнении 4-х групповых средних по признаку PR78:

Группа	N	Среднее	F-критерий	p
1 СТЕПЕНЬ	75	28,257	4,784	0,0035
2 СТЕПЕНЬ	56	30,356		
3 СТЕПЕНЬ	49	31,246		
4 СТЕПЕНЬ	106	29,536		

Kruskal-Wallis Test: $\chi^2 = 14,919$; DF = 3; $p = 0,0019$

Median One-Way Analysis: $\chi^2 = 16,8497$; DF = 3; $p = 0,0008$

Van der Waerden One-Way Analysis: $\chi^2 = 12,3399$; DF = 3; $p = 0,0063$

Сравнение групповых дисперсий:

Siegel-Tukey One-Way Analysis: $\chi^2 = 7,0177$; DF = 3; $p = 0,0713$

11. Если в исследовании производилась проверка идентичности законов распределения количественного признака для двух выборок, то следует подробно описать следующую информацию: 1) объёмы наблюдений в каждой группе сравнения; 2) название статистического критерия согласия; 3) сравнение получен-

ного значения используемого критерия и критического значения. По этим результатам делается вывод о равенстве или неравенстве анализируемых распределений. Рекомендуется для данных проверок использовать такие известные статистические критерии, как двухвыборочный критерий Колмогорова-Смирнова, критерий χ^2 , критерий Андерсона, критерий Ватсона, критерий Купера и т.д.

12. Сравнение групповых параметров рационально всегда производить с проверкой равенства генеральных (популяционных) дисперсий с помощью критериев Фишера, Бартлетта, Романовского, Кохрана, Неймана-Пирсона, Хартли и т.д. Причём для признаков, не имеющих нормальное распределение, использовать непараметрические критерии, например, критерии Ансари-Бредли, Муда, Сиджела-Тьюки, Кейпена, Клотца и т.д. Сообщите название поправки параметров распределения количественного признака при сравнении групп количеством более 2. Если производится проверка равенства групповых параметров сразу в группах количеством более 2, т.е. сравнение 3-х, 4-х и более групп, то после отмены нулевой гипотезы о равенстве всех этих групповых параметров, проводите нужные попарные сравнения, чтобы установить, в каких парах групп эти параметры действительно отличаются. Для производства таких попарных сравнений следует использовать соответствующие методы и критерии. Например, метод прямого сравнения (критерий Тьюки), критерий Шеффе, критерий Стьюдента-Ньюмена-Кейлса, критерий Дункана, критерий Линка-Уоллеса и т.д. При этом использовать поправку Бонферрони, которая уменьшает величину критического уровня статистической значимости используемого критерия. Обязательно указывать, был ли используемый тест одно или же двусторонним, а также были ли сравниваемые группы независимыми, независимыми или же иными.

Продолжение следует.