

ОТЗЫВ

на диссертацию [«Полиморбидность и стоматологическое здоровье пациентов старших возрастных групп и лиц, переживших блокаду Ленинграда»](#), представленной на соискание ученой степени доктора медицинских наук Арьевой Г. Т.

Статистическая часть диссертации, которая послужила доказательной базой научных выводов сделанных автором диссертационной работы, не выдерживает даже самой поверхностной критики. Если книга американских авторов Т.А. Ланга и М. Сесик «Как описывать статистику в медицине» является руководством для авторов, помогающим правильно описывать результаты статистических исследований в научных трудах, то диссертацию Арьевой Г.А. можно считать *образцом того, как не надо описывать статистику в медицине.*

Работа пестрит ошибками при использовании статистических терминов, неверным использованием статистических методов, неправильными записями математических формул, избытком непонятным образом вычисленных уровней значимости. Последнее замечание справедливо не только по отношению к рецензируемой работе, но и к большинству медицинских публикаций, в которых описаны результаты применения статистических методов. Авторы считают вполне достаточным для доказательства тех или иных научных предположений, если в статье присутствует запись типа $p < 0,05$, или $p = 0,...$ При этом не упоминают каким образом вычислено это p , какой метод, или критерий был использован, не говоря уже об обосновании применимости тех или иных статистических методов.

Небрежности в тексте начинаются с раздела 2.3.7, в котором перечислено, что выполнено в работе в части статистического анализа.

Автор пишет:

В соответствии с целью и задачами исследования проводили:

– определение типов распределения данных;

Комментарий. Но в работе не удалось найти соответствующей части текста, которая могла бы послужить обоснованию применения тех или иных параметрических, или непараметрических методов.

Автор пишет:

– расчет элементарных статистик (средние значения – M , ошибки средних значений – m , размах разброса данных – $Min \pm max$);

Комментарий. Нет в математической статистике такого обозначения размаха данных как $Min \pm max$

Автор пишет:

– для анализа межгрупповых различий применяли метод *t*-критерий Стьюдента и методы непараметрической статистики;

Комментарий. В диссертации удалось найти табл. 3, 7, 9, 10 в которых приведены результаты сравнения средних, по-видимому, был использован критерий Стьюдента, на стр. 135 и 170 упоминается критерий Манна-Уитни. Но почему в четырех случаях применялся параметрический критерий, а в одном, на стр. 135 – непараметрический критерий Манна-Уитни нет обоснования.

Автор пишет:

– наличие связи между признаками подтверждалось с использованием множественного регрессионного анализа;

Комментарий. Наличие связи подтверждается корреляционным анализом, а множественная регрессия моделирует форму взаимосвязи, посредством создания математической модели

Автор пишет:

– для выявления наиболее значимых параметров использовали факторный анализ.

Комментарий. В работе не удалось найти соответствующего раздела, в котором были бы описаны результаты факторного анализа.

Это были только комментарии написанного автором на стр. 98 текста диссертации.

Перейдем к анализу результатов проведенных автором статистических исследований.

1. В табл. 3 на стр. 104 приведены результаты сравнения средних. По-видимому, автор использовал критерий Стьюдента, хотя, если сравниваются средние величины в более чем двух группах, следует использовать однофакторный дисперсионный анализ. При парном сравнении средних, если судить по уровням значимости **p**, отличие между всеми средними статистически значимо. Но вопрос, откуда такие значения **p**, как они вычислены? Так как в нашем распоряжении нет исходных данных, пересчитаем **p** для факторов 1 и 2, по значениям средних, стандартному отклонению и объему выборки. Записи 3,02+/-0,99; 2,73+/-0,88 (комбинация знаков +/- вероятно, обозначают символ \pm , который доступен в Word на вкладке Вставка → Символ), что среднее фактора *Образование* равно 3,02,

стандартное отклонение равно $0,99 \times \sqrt{34} = 5,77$. Среднее фактора *Характер проф. деятельности* равно 2,73, стандартное отклонение равно $0,88 \times \sqrt{138} = 10,33$. Введем данные значения в модуль *Другие критерии значимости* программы STATISTICA и получим $p = 0,87$, а в таблице 3 диссертации $p = 0,09$. Естественно результат – нет статистически значимого отличия средних указанных факторов. Аналогично можно пересчитать все p в таблице и убедиться в их неверности. Думаю, что этот печальный вывод справедлив и для всех остальных таблиц (7, 9, 10) сравнения средних в диссертационном исследовании. **Но ведь, по неверным таблицам сделаны и неверные интерпретации, следовательно, выводы диссертационной работы в этой части безосновательны!**

Другие критерии значимости: Таблица средн

Печатать результаты после каждого вычисления Отмена

Сравнение двух коэффициентов корреляции

r1: 0,00 N1: 10 r: 1,0000 Односторонний Вычислить

r2: 0,00 N2: 10 Двусторонний

Сравнение двух средних (нормальное распределение)

C1: 3,02 Ст.откл.1: 5,77 N1: 34 p: ,8750 Вычислить

C2: 2,73 Ст.откл.2: 10,33 N2: 138 Односторонний

Среднее выборки 1 в сравнении со средним 2 Двусторонний

Сравнение двух пропорций (долей)

Пр.1: ,500000 N1: 10 p: 1,0000 Односторонний Вычислить

Пр.2: ,500000 N2: 10 Двусторонний

Следует обратить внимание, что среднее значение фактора *Образование* равно 3,02, почти в 2 раза меньше стандартного отклонения равного 5,77. Среднее фактора *Характер проф. деятельности* 2,73, почти в 3 раза меньше стандартного отклонения 10,33. С другой стороны при таких больших разбросах сравнение средних лишено какого-либо смысла, так как приходим к той же нелепости как и средняя зарплата по стране или, хорошо знакомая докторам – средняя температура по больнице. По утверждению Т.А. Ланга (стр. 35), если стандартное отклонение превышает половину среднего значения, то такие данные не являются нормально распределенными, а значит и некорректно применение критерия Стьюдента.

2. В § 3.2 автором приведены результаты частотного анализа. Но какое отношение к частотному анализу имеют диаграммы размаха на рис.11 и 12, остается загадкой. Судя по подрисуночным названиям рис.11 и рис.12, а также фрагмента текста в конце стр. 120:

На гистограмме (рис. 12) иллюстрированы средняя арифметическая/ошибка среднего/стандартное отклонение и минимальные/максимальные значения показателей наличия МС, ожирения, дислипидемии в исследуемых группах,

автор не имеет понятия, для чего нужны диаграммы размаха и, что на них изображено и, как они должны быть подписаны. Диаграммы размаха неправильно названы гистограммой. Автор прав лишь только в той части, что вместо диаграмм размаха уместно было бы изобразить гистограммы частот.

3. По причинам изложенным в п.1, сомнительного происхождения малые уровни значимости p и в §4.2.1, 4.2.2, 4.2.3, 4.3

4. В § 4.4 (начало) написано

«При проведении линейной регрессии выявлена достоверная обратная связь между показателями: GOHAI и CIRS(G) ($r = -0,103$; $p < 0,05$), возраста и GOHAI ($r = -0,162$, $p < 0,002$), возрастом и коморбидностью ($r = 0,681$, $p < 0,0000$) и возрастом и CIRS(G) ($r = 0,588$, $p < 0,000$). Таким образом, подтверждено, что используемые индексные показатели CIRS(G) и GOHAI пригодны в гериатрической практике, показано, что чем выше показатель GOHAI, тем ниже показатель CIRS(G), следовательно, чем меньше у пациента соматических заболеваний и ниже степень их выраженности, тем выше оценивается качество его стоматологического здоровья»

Очень сомнительно, что при столь малых значения коэффициентов корреляции как $-0,103$ и $-0,162$, уровни значимости p были столь малы, как приводит автор $p < 0,05$ и $p < 0,002$. Но даже если бы статистическая значимость коэффициентов корреляции имела место, корреляции $r = -0,103$ и $r = -0,162$ практически нулевые и лишены смысла как говорить о наличии корреляционной взаимосвязи, так и интерпретировать знак (-). И опять же, причём здесь регрессионный анализ, если идет речь о коэффициенте корреляции.

5. На стр. 135 приведены итоги регрессионного анализа и, опять же удручает безграмотная интерпретация результатов. Коэффициент множественной корреляции $R = 0,254$, что означает слабую множественную взаимосвязь предикторов блока дники, возраст и коморбидность с откликом GOHAI. Коэффициент детерминации $R\text{-Square} = 0,064$, это означает, что построенное уравнение регрессии объясняет всего лишь 6,4% изменчивости отклика относительно среднего значения. Для адекватных уравнений регрессии R и $R\text{-Square}$ должны быть близки к 1, в худшем случае $R\text{-Square}$ должно быть больше, чем 0,5, но, ни как не близким к 0. Таки образом вывод автора:

Таким образом, было показано, что величина GONAI зависит от возраста пациента ($B = 0,19$; $t = -4,09$; $p < 0,00000$) и наличия факта перенесённой в детстве блокады Ленинграда ($B = 4,94$; $t = 3,82$; $p < 0,00005$)

не только лишён смысла, а является в принципе неверным и вредным, так как искажает реальное положение вещей!

Дополнительно, хотелось бы заметить автору, что регрессионный анализ является параметрическим методом, предполагает, что как отклик, так и предикторы количественные переменные и имеют распределение близкое к нормальному. Поэтому составление уравнения с бинарным предиктором *факт перенесенной блокады* не допустимо. В этих случаях целесообразно использование метода общие линейные модели, который также реализован в пакете STATISTICA.

6. В § 4.6, 4.7, 4.10 автор неверно интерпретирует результаты при наличии слабых корреляционных связей, также остается непонятным способ вычисления уровней значимости.

7. В § 4.9 как и в других разделах диссертации автор делает основополагающие выводы, сравнивая частоты, записанные в процентах. Но такие выводы приобретают научный смысл, только в том случае, если показана статистическая значимость отличия этих долей, или использованы таблицы сопряженности для оценки значимости взаимосвязи строк и столбцов таблицы.

8. Во всех разделах, где приведены результаты корреляционного анализа, следовало использовать не корреляцию Пирсона (R), а непараметрические коэффициенты корреляции, так как наряду с количественными переменными были использованы и качественные переменные.

9. На стр. 89 неверно записана формула для расчета *Скорости клубочковой фильтрации (СКФ)* по MDRD:

$$\text{СКФ MDRD7 (мл/мин)} = 170 \times (\text{Scr} \times 0,0113) - 0,999 \times \text{ВЗ} 0,176 \times (\text{Sur} \times 2,8) - 0,17 \times \text{АЛ} 0,318$$

По-видимому, должна быть более сложная формула:

$$\text{MDRD 7 equation (MDRD7)} = 170 \times [\text{serum creatinine(mg/dL)}]^{-0.999} \times [\text{age}]^{-0.176} \times [0.762 \text{ if patient is female}] \times [1.18 \text{ if patient is black}] \times [\text{serum urea nitrogen concentration (mg/dL)}]^{-0.170} \times [\text{serum albumin concentration (g/dL)}]^{0.318}$$

Учитывая достаточно близкие значения средних и наличие больших разбросов в данных под большим вопросом способ вычисления уровней значимости p в § 4.11.

Приведённые замечания ставят под сомнение достоверность всех сделанных автором научных выводов, особенно в той части, где утверждается, что блокадники по некоторым аспектам здоровья приближаются к более молодым группам населения.

Халафян Алексан Альбертович, д.т.н.,
профессор кафедры прикладной математики КубГУ,
statlab@kubsu.ru