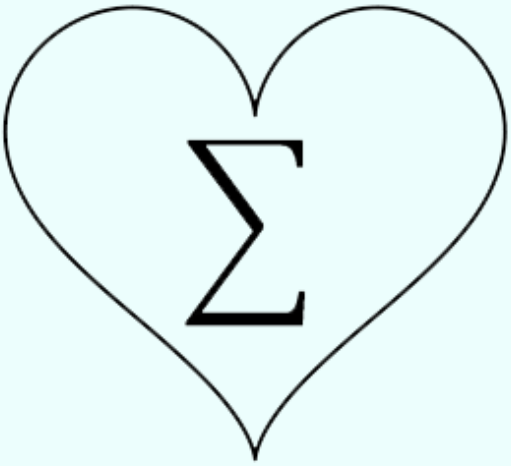
Статистика в кардиологии.

15 лет спустя.

Леонов В.П.



15 лет назад, в 1998 году, в журнале «Кардиология» была опубликована наша статья «Применение методов статистики в кардиологии (по материалам журнала «Кардиология» за 1993–1995 гг.). Показатели смертности от ССЗ в Российской Федерации в целом и по отдельным нозологическим формам более чем в четыре раза превосходят таковую среди населения США, Японии, и Европы. Учитывая актуальность кардиологической проблематики, важно оценить, улучшилось ли за прошедшие 15 лет использование статистики в кардиологии…

БИОМЕТРИКА

http://www.biometrica.tomsk.ru/stat_cardio1.htm

Прочитай и перешли своим коллегам…

Содержание

Введение	с. 2
Что такое ХОРОШО	с. 6
И что такое ПЛОХО	с. 19
Диагноз: патологическая статистика в кардиологии	с. 71
Данные + Статистика = Знания + Экономика	с. 87
Заключение	с. 113
Литература	с. 118

Введение

...Пора чудес прошла, и нам
Подыскивать приходится ***причины***
Всему, что совершается на свете.

_____
В. Шекспир.

15 лет назад, в 1998 году, в журнале «Кардиология» была опубликована наша статья «Применение методов статистики в кардиологии (по материалам журнала «Кардиология» за 1993–1995 гг.) [35]. В этом обзоре были проанализированы 426 статей кардиологической тематики. В заключении обзора говорилось: «Проблема расширения использования современных методов статистической обработки данных в медицинской науке заслуживает отдельного обсуждения. Пока же можно констатировать, что **существующее в настоящее время положение в этой области не отвечает современным возможностям и не способствует максимальному извлечению информации из наблюдений и превращению их в доступное для исследователя знание**». Учитывая острейшую актуальность кардиологической проблематики, а также крайне низкий уровень использования биостатистики в отечественной медицинской науке, важно оценить, улучшилось ли за прошедшие 15 лет данное положение.

В настоящем обзоре представлены результаты анализа 150 статей, опубликованных в издаваемом в НИИ патологии кровообращения им. Е.Н. Мешалкина Минздравсоцразвития России журнале «Патология кровообращения и кардиохирургия» (**ПКК**) в период с 1999 г. по 2013 г. Достаточно подробно описаны результаты анализа около 80 публикаций. Помимо статей из ПКК использованы также несколько публикаций близкой тематики из других журналов.

О том, что в настоящее время проблема качества кардиологических исследований в России стала ещё актуальнее, свидетельствует следующая информация. «Сердечно-сосудистые заболевания (ССЗ) – основные причины смерти населения всех экономически развитых стран мира и стран с переходной экономикой. В последние годы наблюдается эпидемия ССЗ и в некоторых развивающихся странах. По прогнозам экспертов, количество смертей от ССЗ в мире возрастёт за счет увеличения смертности среди мужского населения планеты с 18,1 млн в 2010 году до 24,2 млн в 2030 году. **Показатели смертности от ССЗ в Российской Федерации – одни из самых высоких в мире**. Среди мужского и женского населения

страны в наиболее трудоспособной возрастной группе 25–64 лет **смертность от ССЗ** в целом и по отдельным нозологическим формам **более чем в четыре раза превосходит таковую** среди населения США, Японии, а также экономически развитых стран Европейского региона» [61].

Такое доминирование и рост смертности от ССЗ в России наблюдается на протяжении уже нескольких десятилетий. В большинстве аналитических обзоров основной причиной высокой смертности от ССЗ называется спектр социально-экономических факторов. **Тогда как такая причина, как уровень отечественной медицинской науки, в частности кардиологии, в сравнении с зарубежным уровнем, вообще не рассматривается.** Что, вероятнее всего, объясняется чисто корпоративными интересами. Наглядной иллюстрацией уровня развития российской медицинской науки, в сравнении с мировым уровнем, служит тот факт, что после И.П. Павлова, получившего Нобелевскую премию в 1904 г., ни один отечественный медик или биолог не были отмечены Нобелевской премией по физиологии или медицине [58]. Напомним, что И.И. Мечников, получивший Нобелевскую премию в 1908 г., к тому времени уже много лет жил и работал во Франции в Пастеровском институте. Поэтому тот факт, что «смертность от ССЗ в целом и по отдельным нозологическим формам **более чем в четыре раза** превосходит таковую среди населения США, Японии, а также экономически развитых стран Европейского региона», объясняется ещё и уровнем развития отечественной кардиологии. В частности, весьма низким уровнем использования в кардиологии методов биостатистики. В том, что это различие весьма существенно, можно убедиться, сравнив публикации в российских и зарубежных кардиологических журналах.

Более низкий уровень российской медицинской науки, в сравнении с западной наукой, объясняется не только недостаточным финансированием со стороны государства, но и рядом других причин. В частности, отсталостью организации отечественной медицинской науки, которая, в свою очередь, определяется отношением к науке в целом руководства государства. Свидетельством этого могут служить как наблюдаемые в настоящее время недостаточно эффективные реформы Академии наук, ВАК и технологии присуждения учёных степеней, так и масса других фактов. В качестве примера такого отношения Минздрава к этой проблеме можно привести следующий пример. 10-11 октября 2013 г. в Москве проходила очередная конференция «Информационные технологии в медицине». По адресу http://itm.consef.ru/main.mhtml?Part=30&PubID=662 читаем:

«Отменён Круглый стол по медицинской статистике 10.10.2013. в рамках конференции ИТМ 2013.
Уважаемые коллеги!
Информируем Вас, что круглый стол, запланированный на 10 октября 2013 г. по теме:
«Формирование качественной статистической отчетности. Информационное обеспечение сбора, обработки и анализа статистических данных в здравоохранении»
ОТМЕНЁН.
Указанное решение принято исполнительной дирекцией Оргкомитета конференции **по рекомендации Департамента анализа, прогноза и инновационного развития здравоохранения Министерства** здравоохранения Российской Федерации, полученной во второй половине дня 8 октября 2013 г.

Прочитай и перешли своим коллегам…

В остальном программа XIV Ежегодной специализированной конференции и выставки «Информационные технологии в медицине» сохранилась без изменений.
Приносим извинения за доставленные неудобства.
Исполнительный директор Оргкомитета
Ю. Ю. Мухин»

Кто же рекомендовал отменить обсуждение на этой конференции проблемы обработки и анализа статистических данных в здравоохранении? По адресу https://www.rosminzdrav.ru/ministry/61/2 читаем: «Департамент анализа, прогноза и инновационного развития здравоохранения. Директор Департамента - Семёнов Николай Сергеевич». Отмена "Круглого стола по медицинской статистике" и стала очередной "инновацией развития здравоохранения" этого Департамента.

Одним из последствий непонимания важности этих проблем органами госуправления, как раз и является неквалифицированный статистический анализ собираемых в кардиологии экспериментальных данных. Результатом же дилетантского выполнения данного анализа является резкое снижение информационной отдачи, результативности при сборе кардиологических данных. Поэтому основной целью данного обзора является не сама констатация недостатков используемых авторами статей методов статистического анализа, а указание направлений, реализация которых позволит существенно повысить экономическую эффективность технологии извлечения полезного знания из собираемых кардиологических экспериментальных данных. Данный аспект прекрасно понимается и экономистами, в том числе, и экономистами НИИ патологии кровообращения им. Е.Н. Мешалкина, чью отличную статью на эту тему мы рассмотрим в конце обзора. Экономическая эффективность результатов статистического анализа собираемых кардиологических данных важна ещё и потому, что себестоимость многих видов медицинских анализов весьма высока. Если сделать поиск в интернете на эту тему, то мы увидим, что стоимости отдельных анализов, в расчёте на одного пациента, колеблются от нескольких сотен до нескольких тысяч рублей.

Помимо статьи «Применение методов статистики в кардиологии (по материалам журнала «Кардиология» за 1993–1995 гг.)» [35] по данному направлению нами было опубликовано немало и других близких по тематике статей и обзоров. Вот список некоторых из этих публикаций.

1. Об использовании прикладной статистики при подготовке диссертационных работ по медицинским и биологическим специальностям / Бюллетень ВАК РФ. – 1997. № 5. – С. 56–61. (http://www.biometrica.tomsk.ru/leonov_vak.htm)
2. Долгое прощание с лысенковщиной. (http://www.biometrica.tomsk.ru/lis.htm)
3. Применение статистики в статьях и диссертациях по медицине и биологии. ч. 1. описание методов статистического анализа в статьях и диссертациях /Международный журнал медицинской практики. – 1998. – № 4. – С. 7–12. (http://www.biometrica.tomsk.ru/leonov_1998.htm)
4. Применение статистики в статьях и диссертациях по медицине и биологии. ч. 2. История биометрики и ее применения в России / Международный журнал медицинской практики. – 1999. – № 4. – С. 7–19. (http://www.biometrica.tomsk.ru/history.htm)
5. Применение статистики в статьях и диссертациях по медицине и биологии. ч. 3. Проблемы взаимодействия автор – редакция – читатель / Международный журнал медицинской практики. – 1999. – № 12. – С. 7–13. (http://www.biometrica.tomsk.ru/problem1.htm)

Прочитай и перешли своим коллегам…

6. Применение статистики в статьях и диссертациях по медицине и биологии. ч. 4. Наукометрия статистической парадигмы экспериментальной биомедицины / Международный журнал медицинской практики. – 2002. – № 3. – С. 6–10. (http://www.biometrica.tomsk.ru/part_4.htm)
7. Наукометрика статистической парадигмы экспериментальной биомедицины (По материалам публикаций). Вестник Томского государственного университета. Серия "Математика. Кибернетика. Информатика". №275. Апрель 2002, стр. 17-24. (http://www.biometrica.tomsk.ru/paradigma.htm)
8. Доказательная или сомнительная? Медицинская наука Кузбасса: статистические аспекты. (http://www.biometrica.tomsk.ru/kuzbass1.htm)

Учитывая высокую смертность населения России от кардиопатологий, повышение качества научных исследований в кардиологии является приоритетной задачей для таких «фабрик здоровья», как НИИ им. акад. Е.Н. Мешалкина. Особенность подобных уникальных НИИ состоит в том, что в них сконцентрированы кадры высокой врачебной и научной квалификации. Благодаря огромному потоку пациентов, в них концентрируются также и огромные массивы ценнейшей информации, собираемой в процессе их лечения. Именно эти две составляющие и являются необходимыми условиями эффективности и производительности научных исследований. Однако только наличие необходимых условий, как известно, не всегда являются также и достаточными условиями для достижения конкретных целей. В одной из последних своих работ, известный учёный Д.С. Саркисов, написал: «Сегодня, накануне XXI столетия, полезно взглянуть на минувшую историю теоретической и практической медицины с тем, чтобы наметить возможные направления её дальнейшего развития. Одним из подходов к решению этой задачи является рассмотрение взаимоотношений анализа и синтеза, то есть того, **насколько гармонично всё это время сочеталось накопление новых фактов с их обобщением**. Это является одним из важных условий плодотворного развития теории медицины и её влияния на практическую деятельность врача» [57].

Структура данного обзора такова. В первом разделе «Что такое ХОРОШО» приведены примеры отличного и хорошего описания использованных методов статистики в статьях ПКК. Далее, во втором разделе «И что такое ПЛОХО», приводятся образцы описаний, которые по уровню своего содержания не могут быть включены в 1-й раздел. Т.е. образцы низкого качества. Следует отметить, что во многих описаниях статистических аспектов содержались как положительные, так и отрицательные составляющие. В зависимости от соотношения этих составляющих, данные описания перемещались либо в 1-й, либо во 2-й разделы обзора.

В 3-м разделе «Диагноз: патологическая статистика в кардиологии» приведено обобщение выявленных недостатков использования и описания методов статистики в публикациях ПКК, а также рассмотрены редакционные требования к рукописям статей, публикуемых в кардиологических журналах.

В последнем, 4-м разделе «Данные + Статистика = Знания + Экономика», рассмотрен ряд основных причин, приводящих к недостаточному качеству статистического анализа экспериментальных данных, получаемых кардиологами. Повышение качества данного анализа есть действенное средство роста экономической эффективности деятельности кардиологических

«фабрик здоровья», о чём весьма профессионально написано в анализируемой ниже статье экономистов, работающих в НИИ им. акад. Е.Н. Мешалкина.

Учитывая ряд аспектов, первоначально данный обзор рассылается только ряду руководящих сотрудников НИИ им. акад. Е.Н. Мешалкина. В дальнейшем, после согласования формата публикации краткой версии этого обзора в журнале «Патология кровообращения и кардиохирургия», он будет размещён на сайте БИОМЕТРИКА, а также отправлен в Минздрав, РАН, РАМН, и в некоторые другие организации и сайты. Учитывая пожелание руководителя Центра хирургии аорты, коронарных и периферических артерий, д.м.н., профессора, Заслуженного деятеля науки Российской Федерации А.М. Чернявского, с которым мы в переписке обсуждали этот вариант, в эти организации обзор будет направлен без указания конкретных фамилий авторов анализируемых публикаций.

Что такое ХОРОШО

Умного на свете много,
мало — хорошего.

Александр Солженицын.

Среди всех проанализированных нами статей журнала ПКК за 2010–2013 гг. мы обнаружили следующие статьи, в которых описание использованных методов статистики приведено достаточно полно, корректно и иногда даже образцово. Электронные версии номеров журнала ПКК с 2010 по 2013 гг. доступны читателям по адресу
http://www.meshalkin.ru/nauchnaya_deyatelnost/journal/arhiv_zhurnala.html

1) ПКК, № 2, 2010, с. 13-17. Гендерные особенности приобретённых пороков сердца у больных старшего возраста.
Вот как описывают авторы статистику в разделе «Материал и методы». «Выполнен ретроспективный анализ результатов клинического обследования 1152 пациентов, оперированных по поводу ППС в клинике ФГУ «ННИИПК им. акад. Е.Н. Мешалкина». … Проведён статистический анализ полученных результатов. Количественные величины были представлены как среднее ± стандартное отклонение. Для проверки статистических гипотез о виде распределения был применён критерий Shapiro-Wilk's W. Во всех случаях распределение признаков не соответствовало закону нормального распределения[1] . При выполнении **основной задачи** сравнения двух независимых групп по одному признаку были использованы методы непараметрической статистики (точный критерий Фишера, классический критерий χ^2 по Пирсону, χ^2 с поправкой Йетса). При сравнении относительных частот в двух группах применяли процедуру «Различие между двумя пропорциями». Величину уровня значимости p принимали равной 0,05, что соответствует критериям, принятым в медико-биологических исследованиях. Если значение p было меньше 0,001, то p указывали в формате p<0,001».

Итак, почему данное описание заслуживает положительной оценки? Во-первых, авторы конкретизируют выражения, содержащие 2 числа, соединённых знаком « ± ». Увы, но такая

[1] **NB! Обратите внимание на тот факт, что ВО ВСЕХ СЛУЧАЯХ не наблюдалось нормального распределения.**
Прочитай и перешли своим коллегам…

конкретизация в статьях ПКК встречается не столь уж часто. Далее, авторы сообщают, каким конкретно статистическим критерием проводилась проверка нормальности распределения для количественных признаков. Констатируя при этом отвержение гипотез нормальности для всех количественных признаков. Также авторы перечисляют использованные ими статистические критерии для анализа качественных признаков. Указали авторы и величину критического уровня значимости. Единственный недостаток этого описания заключается в том, что авторы не указали, каким статистическим критерием они пользовались при сравнении групповых средних количественных признаков. Тогда как в тексте статьи результаты таких сравнений приводятся. Учитывая отсутствие согласия с нормальным распределением для этих признаков, авторы, видимо, использовали какой-то непараметрический критерий.

2) ПКК, № 2, 2010, с. 33-37. Раннее выявление нарушений углеводного метаболизма – важный маркёр отдалённого прогноза при инфаркте миокарда.

В разделе «Материал и методы» авторы пишут: «Полученные данные обработаны с использованием программы Statistica 6.0. Средние значения представлены в виде медианы и квартильного отклонения. При сравнении независимых групп использован метод Краскелла – Уоллиса. Взаимозависимость признаков определена методом Спирмена. Время до наступления изучаемого признака определено с помощью регрессионной модели Кокса. Различия считали достоверными при $p<0,05$». Как видим, описание достаточно ясное и понятное. К положительным моментам данной публикации следует также тот факт, что вместо выражений вида «$p < 0,05$» и «$p > 0,05$» авторы везде приводят конкретные значения достигнутого уровня значимости. Положительным моментом публикации следует также считать и наличие в статье таблицы с результатами использования модели Кокса.

Недостатком описания является использование выражений со словом «достоверность». Детальный анализ этой некорректности дан в статье Н.А. Зорина «О НЕПРАВИЛЬНОМ УПОТРЕБЛЕНИИ ТЕРМИНА "ДОСТОВЕРНОСТЬ" В РОССИЙСКИХ НАУЧНЫХ ПСИХИАТРИЧЕСКИХ И ОБЩЕМЕДИЦИНСКИХ СТАТЬЯХ»[24]. Учитывая наличие 3-х групп сравнения, а также достаточно большого количества количественных и качественных признаков, целесообразно использовать в данном исследовании многомерные методы анализа, в частности метод логистической регрессии [44]. Кроме того, при сравнении трёх групп необходимо учитывать проблему множественных сравнений.

3) ПКК, № 2, 2010, с. 52-56. Возможность использования модифицированной шкалы EUROSCORE для оценки годового прогноза коронарного шунтирования у пациентов с мультифокальным атеросклерозом.

В разделе «Материал и методы» авторы весьма детально описывают использованные в исследовании методы статистики. «Статистическая обработка результатов исследования осуществлялась с помощью пакета программ Statistica 8.0 (США). Полученные данные представлены в виде медианы и интерквартильного размаха (25-й и 75-й процентили), средней величины и доверительного интервала. Две независимые группы сравнивались с помощью U-критерия Манна-Уитни, три и более с помощью рангового анализа вариаций по Краскелу-Уоллису с последующим парным сравнением групп тестом Манна-Уитни с применением поправки

Бонферрони при оценке значения p. Анализ различия частот в двух независимых группах проводился при помощи точного критерия Фишера с двусторонней доверительной вероятностью, критерия χ^2 с поправкой Йетса. Анализ прогностических моделей провели с помощью линейной пошаговой регрессии в программе SPSS Statistics 17.0 (США). В последующем рассчитывали скорригированный коэффициент детерминации, показывающий долю объясняемой зависимости. Для определения диагностической ценности прогностической модели использовалась ROC-кривая с последующим определением площади под ней. По данным литературы, диагностически значимым является показатель, превышающий 0,70 [14]». Такое описание может служить примером того, как следует описывать применяемые в исследовании методы статистики. К недостаткам данной статьи следует отнести использование авторами дробных значений баллов, что является некорректным, а также выражений с использованием слова «достоверно».

4) ПКК, № 3, 2010, с. 61-66. Неврологические и нейрокогнитивные функции у пациентов, оперированных по поводу приобретённых пороков сердца.

В отличие от других публикаций данного номера, содержащего материалы Международного конгресса «Сердце – Мозг», в этой статье авторы используют вместо оборотов типа «различие недостоверно» правильные выражения вида «различие не было статистически значимым». Недостатком статистического анализа в данном исследовании следует считать использование t-критерия Стьюдента для дискретной величины, которая априори не может иметь нормального распределения. И в этом случае необходимо было использовать для сравнения групп непараметрические критерии. Учитывая тот факт, что сравнение двух групп проводилось с применением около 20 дискретных признаков, в данном исследовании целесообразно было использовать многомерные методы, например, логистическую регрессию [44] и анализ соответствий.

5) ПКК, № 1, 2011, с. 30-34. Влияние метилпреднизолона на эндотелиальную функцию при кардиохирургических операциях в условиях искусственного кровообращения.

«Статистическая обработка проводилась при помощи программного обеспечения Statistica version 6.0. Для оценки характера распределения в совокупности по выборочным данным использовали тесты Лиллиефорса и Колмогорова – Смирнова. Сравнения двух групп из совокупностей с нормальным распределением проводили с помощью t-критерия Стьюдента для двух зависимых или двух независимых выборок. Для анализа выборочных данных из совокупностей, отличающихся от нормального распределения, использовали непараметрические методы. Для сравнения двух групп применяли критерий Колмогорова – Смирнова. Анализ выборок при повторных измерениях проводили с помощью критерия Фридмана (χ_r^2). Для анализа зависимости количественных признаков выборочных данных из совокупностей с нормальным распределением или без него применяли ранговый коэффициент корреляции Спирмена (r_s). Статистически значимыми считались различия данных и корреляция между данными при $p<0,05$». Отличное описание! К недостатку этого исследования следует отнести отсутствие проверки второго ограничивающего условия (равенства генеральных дисперсий) для использования t-критерия Стьюдента.

6) ПКК, № 4, 2010, с. 46-50. Эффективность иммунохроматографического метода определения сердечного белка, связывающего жирные кислоты, при ранней дифференциальной диагностике острого коронарного синдрома.

Вот как описывают статистические аспекты исследования авторы данной статьи: «Обработка результатов исследования проводилась с помощью программы SPSS 11.05. Определялся характер распределения количественных признаков методом Колмогорова – Смирнова. При отсутствии нормального распределения вычисляли медианы и квартильные интервалы (25 и 75%) и сравнивали независимые выборки с использованием U-критерия Манна – Уитни. Анализ диагностической ценности экспресс-теста сБСЖК при ОИМ выполнялся с помощью ROC-анализа и по общепринятым в медицинской статистике формулам. Критический уровень значимости нулевой статистической гипотезы принимался равным 0,05». Как видим, описание достаточно детальное. Первое замечание касается фразы «… по общепринятым в медицинской статистике формулам». Непонятно, о каких конкретно формулах пишут авторы. В данном случае целесообразно было дать хотя бы ссылки на литературные источники с этими формулами. Далее, поскольку авторы упоминают об использовании ROC-анализа, то в этом случае целесообразно было привести график с ROC-кривой. Кроме того авторы приводят в статье выражения вида «средний возраст 65,20 ± 1,22 лет», но при этом не конкретизируют, что за параметр размещается после знака « ± ». Кроме того, авторы не уточнили, использовалась ли ими поправка Лиллиефорса при проверке нормальности.

7) ПКК, № 3, 2011, с. 23-28. Результаты применения покрытых баллонных катетеров у больных с рестенозом ранее имплантированных коронарных стентов.

В разделе «Материал и методы авторы пишут: «Статистический анализ проводили при помощи лицензионного пакета программ Statistica 6.0. Числовые данные представлены в виде «среднее значение ± стандартное отклонение», логические данные в виде доли от общего числа наблюдений. Для сравнения количественных признаков данных был использован t-критерий Стьюдента. При анализе влияния таких факторов как класс рестеноза и степень остаточного стеноза на частоту развития повторного рестеноза использовался критерий χ^2 Пирсона. Результаты статистического анализа считались значимыми при $p<0,05$». Положительный момент данного описания заключается в конкретизации структуры выражений вида 58,6±8,02, а также в указании критического уровня статистической значимости равного 5%. Кроме того положительным следует считать факт наличия достигнутых уровней статистической значимости в табл. 1 и 2. Однако результаты сравнения групповых средних с помощью критерия Стьюдента могут быть сомнительными по причине отсутствия проверки двух ограничительных условий его применения. В тексте используются некорректные обороты со словом «достоверно». С учётом использования набора количественных и качественных признаков целесообразно было бы использовать для сравнения двух групп метода логистической регрессии, что позволило бы проранжировать анализируемые признаки по интенсивности их различия.

8) ПКК, № 3, 2011, с. 29-34. Причины рестеноза в стенте после интервенционного лечения пациентов с острым коронарным синдромом с элевацией сегмента ST.

В разделе «Материал и методы» авторы пишут: «Статистический анализ осуществляли с помощью пакета прикладных программ Statistica 6.0. Проводилась проверка нормальности распределения количественных признаков, для описания признаков с нормальным распределением использовали среднее с указанием стандартного отклонения, для признаков с отличным от нормального распределения указывали медиану с указанием межквартильного размаха – 25-й и 75-й процентили. Сравнение количественных признаков проводили по критерию Манна – Уитни, сравнение качественных – с использованием таблиц сопряжённости 2 × 2 по критерию χ^2 Пирсона с поправкой Йетса и точному критерию Фишера. Для оценки влияния признака применялся однофакторный анализ с определением ОШ и 95% доверительного интервала (ОШ (95%ДИ)), для многомерной оценки прогностической значимости признаков – дискриминантный пошаговый анализ с включением признаков. Различия принимались как статистически значимые при $p<0,05$». Приведённое выше описание статистических аспектов исследования можно считать образцовым, поскольку в нём подробно описываются детали использования различных статистических методов. Следует также отметить и наличие в тексте статьи достаточно полно описанных результатов, полученных с использованием перечисленных выше методов анализа. В частности, достаточно полно описаны результаты, полученные с помощью пошагового дискриминантного анализа. Однако авторы не уточнили, какой критерий применяли для проверки нормальности.

С учётом использованного набора количественных и качественных признаков, в данном исследовании целесообразно было бы использовать для сравнения групп не только дискриминантный анализ, но и метод логистической регрессии [44], что позволило бы проранжировать анализируемые признаки по интенсивности их различия.

9) ПКК, № 2, 2011, с. 43-48. Клиническая значимость и распространённость «неизменённых» коронарных артерий у больных с острым коронарным синдромом.

В разделе «Материал и методы» авторы пишут: «Статистическая обработка результатов исследования осуществлялась с помощью стандартного пакета статистических программ Statistica for Windows 6.0 (Stat Soft, USA). Использовались в основном непараметрические критерии сравнения выборок (U-критерий Манна – Уитни). Результаты представлены как M (медиана) ± s (стандартное отклонение)». Данное описание является достаточно полным и ясным, и не содержит никаких некорректностей. Более того, данная статья является одной из немногих, в которых авторы не используют архаичные выражения вида «$p<0,05$» и «$p>0,05$», а приводят конкретные величины достигнутого уровня статистической значимости, полученные при проверке тех или

иных статистических гипотез. Ниже приводится таблица 2 с такими результатами.

Таблица 2
Сравнительная характеристика пациентов с ОКС в зависимости от наличия коронарного атеросклероза

Показатели	ОКС с пораженными сосудами (n = 837)	ОКС с «непоражен-ными» сосудами (n = 76)	p
Женщины	260 (31,1%)	33 (43,5%)	0,1
Артериальная гипертензия	658 (78,6%)	76 (100%)	0,1
Сахарный диабет	385 (45,99%)	5 (6,6%)	0,000
Дислипидемия	663 (79,2%)	12 (15,8%)	0,000
Курящие	631 (75,4%)	67 (86%)	0,4
Постинфарктный кардиосклероз	117 (13,97%)	7 (9,2 %)	0,3
Предшествующие чрескожные коро-нарные вмешательства (ЧКВ)	212 (25,3%)	5 (6,6 %)	0,003
Предшествующее коронарное шунтирование (КШ)	55 (6,6%)	0 (0%)	0,04
Периферический атеросклероз	125 (14,9%)	2 (2,6%)	0,01
Острое нарушение мозгового кровообращения (ОНМК)	44 (5,3%)	4 (5,3%)	0,7
Гастроэнтерологическая патология (хроничес-кий гастродуоденит, язвенная болезнь, хрони-ческий холецистит, хронический панкреатит, гас-троэзофагеальная рефлюксная болезнь	115 (13,7%)	19 (25%)	0,04

Учитывая достаточно большой объём наблюдений (n=913), важность и актуальность цели данного исследования, а также количество анализируемых показателей, целесообразно проведение реанализа данного массива с использованием ряда многомерных статистических методов, в частности, многомерных таблиц сопряжённости, лог-линейного анализа, метода логистической регрессии [44].

10) ПКК, № 2, 2011, с. 49-54. Сравнительная оценка профилактического использования внутриаортальной баллонной контрпульсации и левосимендана у больных ишемической болезнью сердца с низкой фракцией выброса левого желудочка.

В разделе «Материал и методы» авторы пишут: «Статистический анализ полученных результатов проводился при помощи таблиц EXCEL, программы STATISTICA 6.1. Для оценки характера распределения в совокупности по выборочным данным использовали тест Колмогорова – Смирнова. Результаты представлены как среднее и стандартное отклонение (M±SD); медиана, 25-й и 75-й процентили. Статистически значимыми считали различия при $p<0,05$. Данные из совокупностей с нормальным распределением сравнивались с помощью t - критерия Стьюдента для независимых выборок. Сравнение данных из совокупностей с распределением, отличающимся от нормального, проводилось с применением критерия Манна – Уитни». Данное описание можно считать образцовым, поскольку оно содержит всю необходимую для читателя информацию. Особо следует отметить использование в этом описании выражения «Статистически значимыми считали различия при $p<0,05$». Единственное упущение в данном исследовании заключается в отсутствии проверки второго условия корректности использования t-критерия Стьюдента. Далее, в тексте используются некорректные обороты с использованием слова «достоверно». Кроме того, авторы не уточнили, использовалась ли ими поправка Лиллиефорса при проверке нормальности.

11) ПКК, № 2, 2011, с. 17-20. Проблемы и возможности эффективной коррекции сердечно-сосудистой недостаточности при ишемической болезни сердца.

В разделе «Материал и методы» авторы пишут: «Статистические расчёты производили на персональном компьютере с использованием программного пакета STATISTICA 6.0. Для оценки показателей до и после лечения применялись параметрические (Стьюдента) и непараметрические критерии (Уилкоксона, Манна – Уитни). Статистический анализ различий между группами больных проводили с использованием однофакторного дисперсионного анализа с последующей оценкой различий между группами с помощью критериев Крускала – Уоллиса, Ньюмена – Кейлса, Данна. **Статистически достоверными** считали различия при $p<0,05$». Приведённое выше описание использованных методов статистики достаточно полное и понятное. К недостаткам этого описания следует отнести отсутствие проверки двух условий корректности для использования критерия Стьюдента и для однофакторного дисперсионного анализа. Кроме того, некорректным является выражение «**Статистически достоверными** считали различия при $p<0,05$». Другим положительным моментом данной статьи следует считать ясное и полное описание приводимых в таблицах дескриптивных статистик: «… все значения представлены в виде среднего значения; медианы; интерквартильного размаха».

12) ПКК, № 4, 2011, с. 49-52. Сравнительная характеристика гипо- и нормотермического искусственного кровообращения и способов защиты миокарда при хирургическом лечении первичных опухолей сердца.

В разделе «Материал и методы» авторы пишут: «Для статистической обработки данных применялся пакет программ статистической обработки данных STATISTICA 6.0. Анализ соответствия вида распределения признака закону нормального распределения проводился с использованием теста Колмогорова – Смирнова. При сравнении групп по количественному признаку использовался непараметрический критерий Манна – Уитни и Колмогорова – Смирнова. Сравнение качественных бинарных признаков проводилось с использованием двустороннего критерия Фишера. За уровень статистической значимости был взят $p<0,05$». Отметим, что в тексте статьи авторы не употребляют выражений со словом «достоверно», а используют правильные выражения вида «статистически значимые различия». Положительным моментом данной публикации следует также считать тот факт, что в таблицах с результатами сравнения групп авторы приводят практически везде не выражения вида «$p > 0,05$» или « $p < 0,05$», а фактические, достигнутые величины статистической значимости для использованных в анализе статистических критериев. Кроме того, авторы не уточнили, использовалась ли ими поправка Лиллиефорса при проверке нормальности.

13) ПКК, № 4, 2011, с. 57-60. Биохимические маркеры постинфарктного заместительного фиброза у мужчин различного возраста.

В разделе «Материал и методы» авторы пишут: «Результаты исследования были проверены на нормальность распределения с использованием критерия Колмогорова – Смирнова. В случае нормального распределения значений применяли параметрический t-критерий Стьюдента. Для значений, закон распределения которых отличался от нормального, был применён непараметрический U-критерий Манна – Уитни. При проведении парных сравнений уровней показателей внутри групп по трём фазам РФ использовали парный критерий Вилкоксона. Данные представлены как средняя арифметическая величина и стандартная ошибка средней ($M \pm m$).

Прочитай и перешли своим коллегам…

Различия считали значимыми при p<0,05». К положительным моментам данной публикации следует отнести проверку нормальности с помощью критерия Колмогорова-Смирнова, использование критерия Манна-Уитни для признаков, не имеющих нормального закона распределения, а также конкретизация величин в выражении M ± m и указание критического уровня статистической значимости. К положительным моментам следует также отнести тот факт, что авторы указывают в тексте статьи конкретные значения достигнутого уровня статистической значимости, вместо использования выражений вида «p<0,05» или «p>0,05». Недостатком проведённого анализа является отсутствие проверки второго ограничения при использовании критерия Стьюдента. Поскольку авторы проводили сравнение между собой более 2 групп, то необходимо было также учесть проблему множественных сравнений, чего не было сделано. Кроме того, авторы не уточнили, использовалась ли ими поправка Лиллиефорса при проверке нормальности.

14) ПКК, № 4, 2011, с. 57-60. Биохимические маркеры постинфарктного заместительного фиброза у мужчин различного возраста.

В разделе «Материал и методы» авторы пишут: «Результаты исследования были проверены на нормальность распределения с использованием критерия Колмогорова – Смирнова. В случае нормального распределения значений применяли параметрический t-критерий Стьюдента. Для значений, закон распределения которых отличался от нормального, был применён непараметрический U-критерий Манна – Уитни. При проведении парных сравнений уровней показателей внутри групп по трём фазам РФ использовали парный критерий Вилкоксона. Данные представлены как средняя арифметическая величина и стандартная ошибка средней (M ± m). Различия считали значимыми при p<0,05». В целом описание статистических аспектов приведено неплохо. За исключением того, что для использования параметрического t-критерия Стьюдента недостаточно проверки лишь одного ограничительного условия – нормальности распределения в обеих группах сравнения. Необходимо выполнение и второго ограничительного условия – равенства генеральных дисперсий. Отметим, что в тексте статьи авторы везде приводят фактические значения достигнутых уровней значимости. Однако авторы не уточнили, использовалась ли ими поправка Лиллиефорса при проверке гипотезы нормальности.

15) ПКК, № 1, 2012, с. 43-46. Гибридные оперативные вмешательства у пациентов с хронической ишемией нижних конечностей.

«Результаты исследования представлены как среднее и стандартное отклонение (M±σ). Для оценки актуарной свободы от тромбоза шунтов использовался метод множественных оценок Каплана – Мейера». В описании авторы конкретизируют смысл выражений вида M±σ, а также используемый статистический метод.

16) ПКК, № 1, 2012, с. 71-74. Вмешательства на межпредсердной перегородке в условиях искусственного кровообращения без окклюзии аорты.

«Статистическая обработка полученных данных проводилась с применением пакета программ Statistica 6.0 с использованием непараметрического критерия Манна – Уитни. Критический уровень значимости p принимался равным 0,05. Результаты представлены как Me (max/min), где Me – медиана, max/min – доверительный интервал». В описании кратко и понятно изложены

статистические аспекты исследования. Единственная некорректность заключается в использовании оборотов типа «достоверная разница».

17) ПКК, № 2, 2012, с. 9-14. Результаты протезирования трикуспидального клапана различными видами протезов у детей.

«Статистическую обработку полученных результатов производили с помощью программы Statistica 7.0. Статистический анализ вели с помощью непараметрических и параметрических методов. Для анализа категоризованных переменных использовали таблицы сопряженности и χ^2. Для оценки выживаемости и свободы от реопераций применяли метод множительных оценок Каплана – Мейера». В тексте статьи авторы конкретизируют использованные статистические методы. «При анализе причин летальности выявлены факторы, статистически значимо с ней связанные: малые возраст (по критериям Вальда – Вольфовица, p = 0,07; Колмогорова – Смирнова, p = 0,01; Манна – Уитни p = 0,017) и вес (по критериям Вальда – Вольфовица p = 0,07; Колмогорова – Смирнова p = 0,025; Манна – Уитни p = 0,004) пациентов. Средний возраст выживших пациентов составил 8,9±4,1 лет, вес – 33,4±12,8 кг, в то время как возраст погибших пациентов составил 4,2±2,9 лет, вес – 19,7±7,6 кг (рис. 1, 2)». Однако авторы не уточняют, какой параметр приводится ими после знака « ± ».

18) ПКК, № 2, 2012, с. 39-43. Состояние микроциркуляции в стенках камер сердца у пациентов пожилого возраста с аортальным стенозом в сочетании с ишемической болезнью сердца.

«Статистическая обработка полученных результатов проводилась с использованием программного пакета Statistica 6.1 (США). Для анализа данных использовали непараметрические методы статистики. Результаты выражали в виде медианы (Me), 25-го и 75-го процентилей (25–75%). Для сравнения двух независимых выборок использовали U-тест Манна – Уитни и критерий χ^2. Для зависимых выборок – тест Вилкоксона. Для всех проведенных анализов различия считали достоверными при уровне значимости $p<0,05$». Описание краткое и полное. Единственный недостаток заключается в использовании некорректного оборота со словом «достоверно».

19) ПКК, № 2, 2012, с. 45-51. Сравнительная клинико-ангиографическая оценка отдалённых результатов коронарного стентирования у больных ишемической болезнью сердца при использовании стентов с различным лекарственным покрытием..

«Статистическую обработку результатов проводили с помощью статистической программы Statistica 6,0. Данные представлены в виде среднее ± стандартное отклонение. Достоверность отличий между группами оценивали при помощи критерия χ^2. Достоверными принимали отличия при $p<0,05$». Недостатком данного описания является отсутствие перечня использованных методов статистического анализа, а также использование оборота «достоверное отличие».

20) ПКК, № 2, 2012, с. 59-63. Функциональное состояние периферического микроциркуляторного кровотока у пациентов с атеросклерозом артерий нижних конечностей.

«Статистическая обработка результатов проведена с помощью программного пакета Statistica 6.0. Данные представлены в виде средних значений и ошибки среднего (M ± m). Для определения различий между группами использовали непарный t-критерий Стьюдента. Для всех проведенных анализов различия считали достоверными при уровне значимости $p<0,05$». Недостатком проведённого анализа и описания является отсутствие проверки условий

корректности использования критерия Стьюдента, а также использование оборота «достоверное отличие».

21) ПКК, № 2, 2012, с. 65-69. Прогноз летальных исходов при диализ-зависимом остром почечном повреждении после кардиохирургических вмешательств.

«Для выявления достоверности различий по анализируемым признакам применялись критерии χ^2 и Манна – Уитни. Для разработки статистической модели использовался логистический регрессионный анализ по многофакторной модели с пошаговым включением переменных. Полученная модель проверялась на состоятельность на экзаменационной группе. Результаты для количественных признаков представлены как медиана, 25-й и 75-й процентили, для качественных как абсолютное количество и процент от общего числа». Данная статья существенно отличается от большинства статей опубликованных в ПКК уровнем использования методов статистического анализа. В отличие от большинства публикаций, в которых авторы ограничиваются лишь сравнением групповых средних, в этой статье авторы использовали современный многомерный метод анализа – логистическую регрессию. Этот факт свидетельствует о понимании ими того, что ценная информация сосредоточена не только в средних величинах сравниваемых подгрупп, но также и в связях между анализируемыми признаками. Отметим как положительный момент данной статьи также и достаточно полное и понятное изложение результатов достаточно сложного метода – логистической регрессии. В частности, авторами приведена таблица с результатами сравнения ROC-кривых для разных моделей. К недостатку этой публикации можно отнести использование авторами оборотов «достоверность различий».

22) ПКК, № 1, 2013, с. 5-9. Процедура Росса у пациентов с выраженной систолической дисфункцией левого желудочка.

«Статистическая обработка полученных результатов осуществлялась с помощью программы «Statistica 6.0». Результаты представлены как среднее и стандартное отклонение (M ± σ), медиана и 25-й, 75-й процентели. Для сравнения двух групп применяли критерий Манна – Уитни. Анализ двух выборок при повторных измерениях проводили с помощью парного теста Уилкоксона. Статистически значимыми считались различия данных при $p<0,05$». (Слово «процентели» следует писать как «процентили»). Отметим также, что математической статистике принято греческими буквами обозначать генеральные параметры распределений, а выборочные параметры – соответствующими латинскими аналогами. Поэтому для обозначения выборочного стандартного отклонения вместо греческой буквы «σ» следует использовать латинскую букву «s».

23) ПКК, № 1, 2013, с. 21-24. Непосредственные и отдаленные результаты различных методов реконструкции каротидной бифуркации.

«Статистическую обработку материала исследования осуществляли с помощью программы «Statistica 6.0» (США). Проверку нормальности распределения параметров проводили с использованием критерия Шапиро – Уилка. Для параметров, отвечающих критериям нормального распределения, применяли параметрические методы статистики, данные представляли в виде среднего арифметического ± стандартное отклонение (M ± σ). В остальных случаях использовали непараметрические методы статистики. Достоверность отличий оценивали по критерию χ^2. При сравнении качественных признаков с числом наблюдений менее 5 применяли точный критерий Фишера. Достоверными принимали отличия при $p<0,05$, когда вероятность различий была больше 95%». Описание достаточно подробное. К недостаткам можно отнести упоминание критерия

Пирсона χ^2 без конкретизации, в каких случаях использовался этот критерий, поскольку он используется при проверке различных статистических гипотез. Кроме того авторы не уточнили, какие именно непараметрические методы ими использовались. Для обозначения выборочного стандартного отклонения вместо греческой буквы «σ» следует использовать латинскую букву «s». Недостатком описания является использование оборота «достоверное отличие».

24) ПКК, № 1, 2013, с. 55-60. Отдалённые исходы каротидной хирургии с позиции мультифокальности атеросклеротического процесса у больных с метаболическими расстройствами.

«Статистический анализ результатов исследования проводили с использованием лицензионной программы SPSS 18.0. Для описания числовых значений выборочных данных при нормальном распределении использовались выборочное среднее и выборочное стандартное отклонения. Количественные признаки с асимметричным распределением описывались с помощью медианы и процентилей. Для сравнения групп по количественным признакам использовались методы параметрической статистики, поскольку распределение признака в группах больных было нормальным, что было установлено с помощью теста Колмогорова – Смирнова ($p<0,05$). Для проверки статистических гипотез для двух групп использовался критерий Стьюдента как частный случай однофакторного дисперсионного анализа. Сравнение качественных признаков проводилось с помощью таблиц сопряжённости (**хи-кватрат** (***так написано в оригинале статьи*** **В.Л**.) по методу Пирсона с поправкой Йетса). Если ожидаемые значения после составления таблиц сопряженности при сопоставлении качественных признаков не превышали 5, то их сравнение проводилось с помощью точного критерия Фишера попарно. Анализ выживаемости проводился путем построения таблиц дожития с графическим изображением кривых выживаемости по методу Каплана – Майера. Сравнение кривых выживаемости проводилось с помощью логрангового критерия и обобщенного критерия Уилкоксона (критерий Гехана). Разница считалась значимой при $p<0,05$».

Данное описание является одним из наиболее ёмких и подробных среди публикаций журнала ПКК. И это можно только приветствовать. В то же время и в этом описании есть отдельные детали, которые требуют уточнения. Например, авторы упоминают о наличии количественных признаков с асимметричным распределением вероятностей. Что говорит об отклонении для них гипотезы нормальности. Однако при этом авторы не уточняют, какой конкретно статистический критерий в этом случае использовали вместо критерия Стьюдента. Не сообщают авторы и о проверке второго условия применимости критерия Стьюдента. Как правильно отмечают авторы статьи, критерий Стьюдента действительно является частным случаем классического однофакторного дисперсионного анализа (ANOVA). Однако, как и для дисперсионного анализа, так и для критерия Стьюдента, необходимо одновременное выполнение двух ограничивающих условий: нормальность распределения во всех группах сравнения, и равенство генеральных дисперсий для всех групп сравнения.

25) ПКК, № 1, 2010, с. 31-34. Динамика диастолической и систолической функции левого желудочка у больных ишемической болезнью сердца с выраженной постинфарктной левожелудочковой дисфункцией после хирургического лечения.

К положительным моментам публикации следует отнести пояснение структуры выражения для средних величин вида 55,3±9, а также факт упоминания критического уровня статистической значимости, и использованных статистических критериев сравнения центральных мер. К недостаткам публикации следует отнести использование оборотов со словом «достоверно».

26) ПКК, № 1, 2010, с. 60-63. Диагностика асинергии миокарда с помощью лимитированной эхокардиографии, проведённой на портативном ультразвуковом аппарате у пациентов с острым коронарным синдромом.

Авторы пишут: «(здесь и далее данные представлены как M ± m)», однако при этом не уточняют, какие конкретно выборочные параметры распределений связаны знаком ± . Такое умолчание было бы вполне объяснимо, если бы редакционные требования прямо указывали бы, что есть параметр M, и что есть параметр m. Далее авторы пишут: «Анализ результатов проводился с использованием статистических пакетов SPSS для Windows. Для оценки информативности диагностического теста применяли четырехпольную таблицу сопряжённости, расчёт показателей информативности диагностического теста (чувствительность, специфичность, отрицательная прогностическая ценность (ОПЦ), положительная прогностическая ценность (ППЦ), прогностическая точность (ПТ), ROC-анализ). Для анализа согласованности результатов двух методов диагностики использовали расчёт коэффициент корреляции по Спирмену, расчёт меры согласия К. За достоверность изучаемых параметров принимали уровень $p<0,05$». Если относительно вида коэффициента корреляции вопросов не возникает, то относительно меры согласия «К», которую использовали авторы, не всё понятно. Можно предположить, что это не что иное, как показатель «альфа Кронбаха». Однако в литературе и в описании пакета SPSS данный показатель обозначается не как «мера К», а именно как «альфа Кронбаха». К недостаткам статьи можно отнести также использование некорректных оборотов со словом «достоверно». Учитывая фиксирование у наблюдаемых пациентов с установленным диагнозом ОКС или без ОКС многих клинических показателей, целесообразно было использовать для достижения поставленных целей исследования пошаговые алгоритмы логистической регрессии [44], и для оценок уравнений на каждом шаге можно было построить ROC-кривые. На основе этих оценок и следовало далее выбрать наиболее оптимальное уравнение, включающее предикторы, получаемые при использовании лимитированной эхокардиографии, проведённой на портативном ультразвуковом аппарате.

27) ПКК, № 2, 2010, с. 29-32. Влияние уровня эктопической активности на результативность хирургического лечения предсердной эктопии.

.

В разделе «Материал и методы» авторы пишут: «Все результаты выражены как арифметическое среднее ± стандартное отклонение. Для оценки достоверности различий между средними и долями при соответствии распределения вариант закону нормального распределения использовали t-критерий Стьюдента при значении $p=0,05$ (уровень достоверности для медиан 95%) и χ^2-критерий для дискретных переменных. Для оценки статистических различий изменений основных характеристик внутри групп и между группами использовали непараметрический Wilcoxon-Mann-Whitney U-тест. Для оценки корреляции между изучаемыми явлениями использовали коэффициент линейной корреляции Pearson». Положительным моментом данного описания можно считать описание выражения вида 21,2±4,2, а также конкретизацию использованных методов статистического анализа. Однако авторы не проверяли второго условия корректного применения критерия Стьюдента. Напомним, что одновременное выполнение обоих условий встречается крайне редко. К недостаткам данной публикации можно отнести также частое использование авторами выражений вида $p < 0,05$ и $p > 0,05$, вместо записи конкретной величины достигнутого уровня значимости. Кроме того, авторы используют некорректные выражения со словом «достоверность». Учитывая наличие двух групп сравнения, которые описывались наборами качественных и количественных признаков, целесообразно было использовать в данном исследовании такой метод, как логистическая регрессия [44].

Прочитай и перешли своим коллегам…

28) ПКК, № 2, 2010, с. 38-41. Содержание простагландина j(2) в крови и величина комплекса интима-медиа сонных артерий у женщин с артериальной гипертензией в постменопаузальном периоде.

В разделе «Материал и методы» авторы пишут: «Статистическую обработку данных проводили с помощью программного пакета STATISTICA 6,0 с оценкой числовых переменных – средней арифметической (M), ошибки средней (m) и определением достоверности различий (p). Достоверность различий полученных результатов оценивали по t-критерию Стьюдента с поправкой Бонферони [2]. Различия считали достоверными при 5% уровне значимости ($p<0,05$)». Положительным моментом данного описания можно считать описание выражения $M \pm m$, а также использование поправки Бонферрони. Недостатком же является использование выражения со словом «достоверность». Кроме того, авторы не проверяют ограничения на использование критерия Стьюдента.

29) ПКК, № 4, 2010, с. 8-11. Динамика восстановления показателей систолической и диастолической функции левого желудочка у детей раннего возраста с коарктацией аорты и сниженной фракцией выброса левого желудочка.

В разделе «Материал и методы» авторы пишут: «Статистическая обработка полученных данных проводилась с помощью программы Statistica 6, с использованием t-теста для двух независимых выборок. Результаты представлены как среднее и стандартное отклонение ($M \pm \sigma$). Статистически значимыми считались различия данных при $p<0,001$». К положительному моменту данного описания следует отнести объяснение выражений вида 3,35±0,73. Однако в математической статистике греческими буквами обозначают генеральные параметры, а выборочные параметры обозначаются их латинскими аналогами. Поэтому вместо выражения $M \pm \sigma$ следовало использовать выражение $M \pm s$. Однако сам факт объяснения, что такое выражение вида 3,35±0,73, является положительным моментом данной публикации. Далее, авторы не пояснили, почему ими был выбрана столь малая величина критического уровня значимости, равная 0,001. Поскольку авторы сравнивали между собой 5 групп, в силу чего для учёта проблемы множественных сравнений они и могли выбрать столь малый уровень значимости. Например, если использовать поправку Бонферрони, то критический уровень значимости следует уменьшить в 5*(5-1)/2=10 раз. Если выбор такого критического уровня значимости был сделан именно по этой причине, то в этом случае нужно было обосновать его.

К недостаткам данной публикации следует отнести тот факт, что авторы использовали t-критерий Стьюдента без проверки ограничений на применение данного критерия. Учитывая тот факт, что при обследовании пациентов фиксировались как качественные, так и количественные признаки, для сравнения групп целесообразно было использовать логистическую регрессию [44].

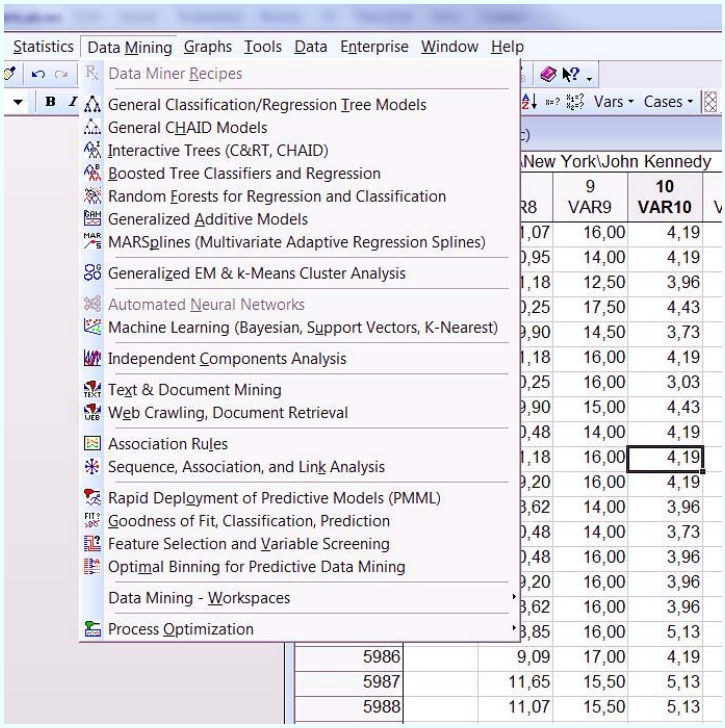
И что такое ПЛОХО

Единственная настоящая ошибка
— не исправлять своих
прошлых ошибок.

Конфуций

1) ПКК № 3, 2007. С. 7–14. Пятилетний опыт использования биологических протезов «КЕМКОР» в митральной позиции.

В статье сообщается: «Для проведения статистической обработки полученных численных значений использовался программный пакет «STATISTICA 6.1 Data Miner», модули описательной статистики и Data Mining, предоставленный компанией StatSoft Russia. **Уровень значимости p принимали больше 0,05**». Далее в разделе «Результаты» авторы



пишут: «Получены результаты прогнозирования вероятности развития тромботических осложнений в отдаленные сроки после изолированного протезирования митрального клапана биопротезом «КемКор» с применением статистической программы Data Mining [4]. Всего было изучено прогностическое значение 90 различных показателей (переменных)». Однако при этом авторы ничего не сообщают о том, какие конкретно методы Data Mining (рус. *добыча данных, интеллектуальный анализ данных, глубинный анализ данных; см.* http://ru.wikipedia.org/wiki/Data_mining) были ими использованы. В то же время в статье приводится порядка 100 процентных соотношений и выражений вида 50,14 ± 9,35. Складывается впечатление, что эти выражения и есть результаты использования интеллектуальных процедур Data Mining. Слева приведено изображение меню Data Mining, из которого видно, что таких процедур довольно много. Отметим, что результаты использования этих процедур содержат весьма большое количество числовых параметров. Однако никаких конкретно числовых параметров, полученных с помощью этих процедур, авторы в своей статье не приводят.

Далее рассмотрим утверждение авторов «Уровень значимости p принимали больше 0,05». Подобное утверждение свидетельствует о том, что авторы статьи имели на момент публикации статьи весьма смутное представление о смысле такой величины, как «уровень значимости». Обратимся к толкованию этого термина в «Статистический словарь»/ Гл. ред. М.А. Королёв, – 2-е изд., перераб. и доп. - М.: Финансы и статистика. – 1989. – 623с. На стр. 542 читаем: «**УРОВЕНЬ ЗНАЧИМОСТИ** – одна из характеристик качества критерия статистической проверки гипотез. Пусть выдвинута гипотеза H_0 (основная, или "нулевая"). **Всякое стат. решение, принимаемое на основе ограниченного ряда наблюдений, неизбежно сопровождается вероятностью ошибочного заключения.** С вероятностью альфа гипотеза H_0 может оказаться отвергнутой, в то время как на самом деле она является справедливой (ошибка первого рода), или, наоборот, с вероятностью бэта может быть принята гипотеза H_0 в то время как на самом деле она является ошибочной (ошибка второго рода). ... В частности, при фиксированном объёме выборки обычно задаются величиной альфа вероятности ошибочного отвержения проверяемой гипотезы H_0. Эту вероятность ошибочного отклонения "нулевой" гипотезы принято называть **УРОВНЕМ ЗНАЧИМОСТИ**... На практике часто пользуются след. стандартными значениями альфа: 0,1 , 0,05 , 0,025 , 0,01 , 0,005 , 0,001. Особенно распространённой является величина **УРОВНЯ ЗНАЧИМОСТИ** альфа равная 0,05. Она означает, что в среднем в пяти случаях из ста ошибочно отвергают высказанную гипотезу при пользовании данным критерием статистическим». Следуя же утверждению авторов статьи, получается, что они соглашались ошибочно отвергать нулевую статистическую гипотезу и с вероятностью большей, нежели 5%. Например, с вероятностью 10%, 25%, 50% и даже 100%! Ведь все эти величины более 5%. Т.е. вместо корректного выражения типа «Критический уровень значимости принимался равным 5%», авторы использовали некорректную, совершенно ошибочную формулировку.

Дополним это толкование термина «**УРОВЕНЬ ЗНАЧИМОСТИ**» обратившись к определению понятия доверительной вероятности. Для этого используем популярную среди биологов и медиков книгу Г.Ф. Лакина "Биометрия" (Москва, изд-во "Высшая школа", 1990. – 352 с.), в которой на с. 107 читаем следующее. «...С доверительной вероятностью тесно связан уровень значимости альфа, под которым понимают разность альфа=1–P». Авторы никак не поясняют, почему «Уровень значимости p принимали больше 0,05». Однако это утверждение бессмысленно ещё и вот по какой причине. В статистическом анализе уровень значимости для используемых статистических критериев принимается **РАВНЫМ** некоторой постоянной величине. Эта величина используется в качестве граничного, критического уровня значимости. К примеру, если критический уровень значимости принимается **РАВНЫМ** 0,05, а для используемого статистического критерия фактический, достигнутый уровень значимости оказывается более чем p=0,05 , например, p=0,12, т.е. 12%, то в этом случае нет оснований отклонить нулевую гипотезу, и она принимается. Нулевая же гипотеза в каждом конкретном случае может быть разная. Например, с помощью t-критерия Стьюдента проверялась гипотеза о равенстве двух генеральных средних в двух совокупностях. Либо нулевая гипотеза заключалась в утверждении равенства двух генеральных коэффициентов корреляций, либо в равенстве двух генеральных дисперсий, либо в равенстве генерального коэффициента корреляции нулю, и т.д. Однако наибольшая некорректность это утверждения заключается в использовании авторами отношения «**больше**».

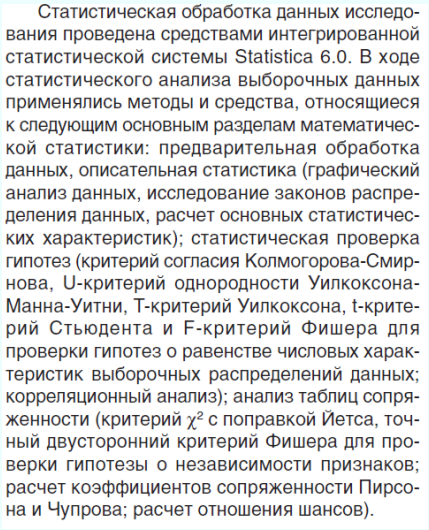
Прочитай и перешли своим коллегам…

Ведь при этом авторы не указывают, до какого конкретного значения, **большего чем 0,05**, они принимают уровень значимости. Поскольку уровень значимости есть вероятность, которая может принимать значения в интервале от 0 (НЕВОЗМОЖНОЕ СОБЫТИЕ), до 1 (ДОСТОВЕРНОЕ СОБЫТИЕ), то утверждение «Уровень значимости p принимали больше 0,05» можно понимать так, что авторы статьи принимали в том числе и такие уровни значимости, как 0,5 ; 0,9; 0,95, и даже 1. Ведь все эти значения тоже больше 0,05!

Если же обратиться к статье «Непосредственные результаты протезирования аортального клапана каркасными биологическими протезами «БиоЛАБ КА/ПТ» (ПКК № 1, 2011. С. 16-20), то мы увидим, что в обеих статьях есть те же авторы. Однако в этой новой статье (её анализ приведён ниже) авторы в разделе «МАТЕРИАЛ И МЕТОДЫ» ничего не говорят о том, какой уровень значимости был ими использован. Однако в тексте статьи встречаются выражения типа «p<0,05». Из чего можно сделать вывод о том, что неявно в данном случае уровень значимости принимался равным 0,05, а не «больше 0,05». Вероятно, что за 4 года, прошедшие с момента публикации первой статьи, данные авторы осознали некорректность выражения «Уровень значимости p принимали больше 0,05», и поэтому в новой статье уже не использовали его.

Бессмыслицы в описании используемого уровня статистической значимости в медицинских публикациях весьма часты. Что объясняется не только недостаточным уровнем знаний медиков в основных понятиях статистики, но также и отсутствием в редакционных коллегиях медицинских и биологических журналов сотрудников, обладающих этими знаниями в необходимом объёме. В качестве примера такой нелепости приведём статью «Анализ взаимосвязи полиморфизма C677T гена метилентетрагидрофолатредуктазы с клиническими проявлениями атеросклероза»; (авторы М.Г. Спиридонова[1], В.А. Степанов[1], В.П. Пузырев[1], Р.С. Карпов[2]. [1] Научно-исследовательский институт медицинской генетики Томского научного центра Сибирского отделения РАМН (директор - академик РАМН В.П. Пузырев). [2] Научно-исследовательский институт кардиологии Томского научного центра Сибирского отделения РАМН (директор - академик РАМН Р.С. Карпов). Журнал «Генетика», вып. 9, 2000, стр. 1269-1273. (URL: http://www.biometrica.tomsk.ru/kk/index_3.htm#33) Цитата из статьи (с. 1270): «Для всех статистических тестов в качестве критерия статистической достоверности **рассматривался уровень значимости более 0,95**». Как видим, помимо того, что два академика РАМН, два директора НИИ РАМН допустили столь грубую ошибку в описании уровня значимости, эту же ошибку допустила и редакция столь известного журнала, как журнал «Генетика». Что вполне объяснимо, поскольку один из авторов этой статьи, академик РАМН, директор НИИ медицинской генетики РАМН В.П. Пузырев, является членом редакционной коллегии данного журнала.

2) ПКК, № 4, 2007, с. 46-50. Сравнительные долгосрочные результаты операций каротидной эдарэктомии с пластикой заплатами из ксеноперикарда, обработанного диэпоксисоединениями, и аутовены.



Статистическая обработка данных исследования проведена средствами интегрированной статистической системы Statistica 6.0. В ходе статистического анализа выборочных данных применялись методы и средства, относящиеся к следующим основным разделам математической статистики: предварительная обработка данных, описательная статистика (графический анализ данных, исследование законов распределения данных, расчет основных статистических характеристик); статистическая проверка гипотез (критерий согласия Колмогорова-Смирнова, U-критерий однородности Уилкоксона-Манна-Уитни, T-критерий Уилкоксона, t-критерий Стьюдента и F-критерий Фишера для проверки гипотез о равенстве числовых характеристик выборочных распределений данных; корреляционный анализ); анализ таблиц сопряженности (критерий χ^2 с поправкой Йетса, точный двусторонний критерий Фишера для проверки гипотезы о независимости признаков; расчет коэффициентов сопряженности Пирсона и Чупрова; расчет отношения шансов).

Слева приведено описание использованных авторами данной статьи методов статистики. Рассмотрим следующее утверждение авторов: «В ходе статистического анализа выборочных данных применялись методы и средства … **для проверки гипотез о равенстве числовых характеристик выборочных распределений данных..**.». Итак, авторы пишут о том, что они применяли методы и средства статистического анализа «… **для проверки гипотез о равенстве числовых характеристик выборочных распределений данных**». Из этого утверждения следует явное непонимание смысла статистических гипотез, о проверке которых пишут авторы статьи. Авторы формулировали свои гипотезу, которые далее проверяли различными методами, для «**числовых характеристик выборочных распределений данных**». Т.е., если ими были вычислены СРЕДНИЙ ВОЗРАСТ для первой группы пациентов (n=21), равный 56 лет, и СРЕДНИЙ ВОЗРАСТ для второй группы пациентов (n=106) равный 57,5 года. То в этом случае утверждение авторов означает, что проверяемые ими гипотезы для этого случая формулируются так: нулевая гипотез H_0: 56 = 57,5 ; альтернативная гипотеза H_1: 56 ≠ 57,5. Но ведь уже из школьного курса арифметики известно, что число 57,5 больше числа 56. И для этого нет нужды использовать никакие методы статистики. В том числе ни U-критерий Уилкоксона-Манна-Уитни, ни T-критерий Уилкоксона, ни t-критерий Стьюдента и т.п. Статистические гипотезы тем и отличаются от арифметических соотношений «равно», «больше–меньше» и «не равно», что они проверяют гипотезы не для соотношений **ВЫБОРОЧНЫХ ХАРАКТЕРИСТИК**! А что же тогда они проверяют, о чём не ведают авторы статьи? Об этом очень подробно написано практически во всех учебниках по статистике… И этому мы учим на наших выездных семинарах по статистике (http://www.biometrica.tomsk.ru/seminar.htm) уже второй десяток лет.

В описании процедуры статистического анализа данных, использованных в данном исследовании, перечисляются 9 разнообразных методов. Это критерий Колмогорова-Смирнова, U-критерий Уилкоксона-Манна-Уитни, T-критерий Уилкоксона, t-критерий Стьюдента, F-критерий Фишера, анализ таблиц сопряжённости с критерием Пирсона χ^2 с поправкой Йетса, точный двусторонний критерий Фишера и т.д. Однако в разделе «Результаты и обсуждение» не ни значений этих критериев, ни достигнутых для них значений величин статистической значимости, нет и коэффициентов сопряжённости Пирсона и Чупрова, нет и отношений шансов. При этом в тексте статьи используется несколько десятков процентных соотношений. В этой связи возникает вполне закономерный вопрос: с какой целью авторы статьи перечислили так много статистических критериев и методов, не приведя при этом никаких конкретных результатов их использования? Напомню, что методологический принцип «бритвы Оккама» гласит: «Не следует

привлекать новые сущности без необходимости». В данном случае налицо нарушение этого принципа. Или всё же необходимость упоминания столь большого количества статистических методов была? Тогда в чём она заключалась? Какова была её цель?

Авторы пишут: «Таким образом, полученные данные показали отсутствие **статистически достоверной разницы** при применении разного вида заплат в оценке долгосрочных результатов операций по таким критериям, как толщина неоинтимы, наличие пристеночного тромба и атеросклеротических изменения сосудистой стенки в месте пластики». Обороты «**достоверно выше**» и «**достоверно ниже**» встречаются также и в других утверждениях в данной статье. Данные обороты некорректны [24], о чем мы уже упоминали выше.

Авторы пишут: «Средний возраст пациентов составил 57±7,3 года». Однако при этом не уточняют, какой выборочный параметр приводится после знака « ± ». Это может быть как стандартное отклонение, так и стандартная ошибка среднего, либо полуширина доверительного интервала. В принципе такие уточнения могут и не потребоваться, если в редакционных требованиях к рукописям статей имеется конкретизирующее требование приводить после этого знака величину конкретного параметра, например, величину выборочной ошибки среднего. Что было бы вполне разумно, поскольку в этом случае заинтересованные читатели могут без особого труда самостоятельно вычислить доверительные интервалы для нужных величин, и сравнить их с доверительными интервалами, полученными по собственным данным. Но ведь этого указания нет в редакционных требованиях журнала ПКК.

3) ПКК, 1999, № 1, с. 69-73. Биопротезы в сердечно-сосудистой хирургии. 20-летний опыт Кемеровского кардиологического центра.

На стр. 69 читаем: «Средний возраст пациентов этой группы - 32,3± 1,2 года». Аналогичные выражения для средних величин встречаются в тексте статьи и далее. Однако нигде при этом не сообщается, что за величины размещены автором после символа ± . Вероятнее всего отсутствие этого пояснения вызвано тем, что в редакционных требованиях журнала ПКК ничего не говорится о том, каким образом описывать результаты статистического анализа.

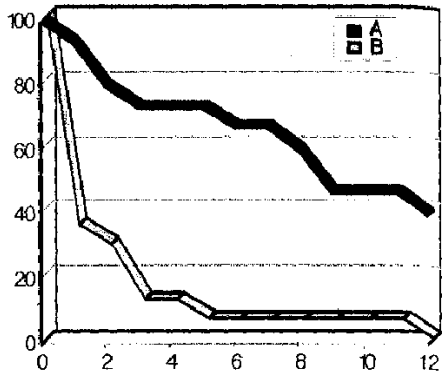


Рис. 1. Актуарное отсутствие тромбозов биопротезов из аллогенных артерий (А) и аллогенных вен (В) в сроки до 12 месяцев после операции (p=0,001)

На стр. 70 автор приводит рис. 1, в подписи под которым приведено значение p=0,001. Видимо, это есть значение достигнутого уровня статистической значимости для некоего статистического критерия. Однако каким методом, и с использование какого конкретно статистического критерия было проведено сравнение двух групп наблюдений, автор не сообщает.

4) ПКК, 2007, № 4, с. 11-15. Динамика показателей липидного метаболизма у больных ишемической болезнью сердца после коронарного шунтирования.

Прочитай и перешли своим коллегам…

На странице 12 авторы сообщают, что «Полученные результаты обрабатывались методом вариационной статистики». Т.е. авторы называют науку статистику ***методом***. Тогда как в действительности данная наука содержит в себе огромное количество самых разных ***методов***. Этот факт подтверждается, в том числе, и использованием во многих публикациях оборота «применялись ***методы*** вариационной статистики». Ниже мы приведём фрагмент из статьи «Эндотелиальная дисфункция и фактор курения у пациентов с ишемической болезнью сердца при тестировании лучевой артерии перед коронарным шунтированием» (ПКК, № 1, 2012, с. 33-37) где авторы как раз используют данный оборот. При обсуждении этого фрагмента отметим также и анахронизм выражения «вариационная статистика».

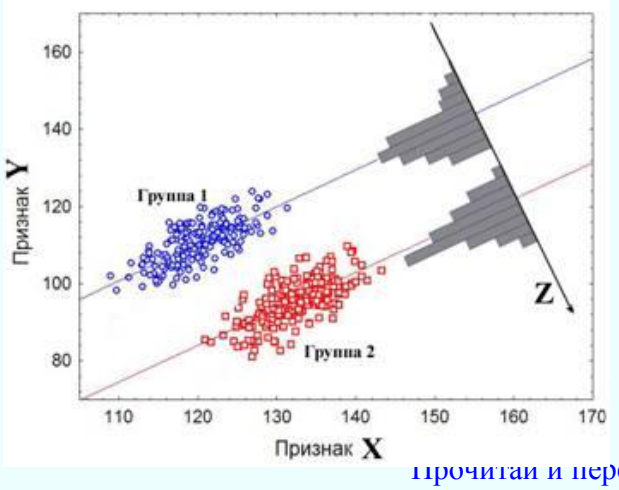
> Полученные результаты обрабатывались методом вариационной статистики с применением t-критерия Стьюдента для средних арифметических (p<0,05).

Таблица 2

Концентрации общего холестерина, фракций и субфракций сывороточных ЛП (по ХС) в венозной крови у больных ИБС (n=16) до и после операции КШ

Фракции и субфракции ЛП, мг/дл	До операции КШ	После операции КШ	Разница %
ЛПОНП	26,4±7,4	18,2±6,7	-31,2
$ЛПОНП_{1-2}$	2,0±0,6	1,6±0,5	-24,1
$ЛПОНП_{3-5}$	24,4±7,0	16,6±6,4	-31,7
ЛППП	21,4±6,2	18,0±8,4	-16,0
ЛПНП	123,7±16,8	128,4±17,1	+3,8
$ЛПНП_{1-3}$	102,7±15,2	110,5±15,5	+7,6
ЛПВП	35,5±4,9	35,0±2,7	-1,3
$ЛПВП_2$	16,7±5,4	16,0±6,2	-4,1
$ЛПВП_3$	18,8±4,3	19,0±5,5	+1,1
Общ. ХС	185,6±21,3	181,6±21,4	-2,1
Общ. ТГ	164,5±25,2	134,0±25,7	-18,5

Далее авторы сообщают, что использовали t-критерий Стьюдента. Однако при этом не проводят проверку ограничений на использование данного критерия. Не обнаруживаются в статье и непосредственные результаты использования t-критерия Стьюдента, например, значения этого критерия при сравнении двух групп, и значения достигнутых уровней значимости для них. В статье сообщается о формировании 3-х групп сравнения. И в этом случае при использовании t-критерия Стьюдента возникает проблема множественных сравнений, что требует изменения критического уровня значимости. Однако и об этом в статье также ничего не сообщается. Из чего можно сделать вывод о том, что авторы статьи не знают об этой специфике использования t-критерия Стьюдента. Далее, как и во многие других статьях, публикуемых в журнале ПКК, авторы используют обороты со словом «достоверно». Что в данном контексте является некорректным [24]. Как и во многих других статьях этого журнала, авторы используют выражения вида 214,0 ± 24,9 , не сообщая при этом, что за величины расположены в них после знака ± . На странице 13 авторы приводят в табл. 2 средние значения концентраций липидов. А на следующей странице приведён график, отражающий отклонение показателей липидного спектра после операции коронарного шунтирования при неосложнённом и осложнённом течении. Основным же недостатком данного исследования является «плоский» подход к изучению динамики изучаемых показателей. Суть этого «плоского» подхода заключается в использовании одних лишь процедур сравнения групповых средних, при полном игнорировании ценнейшей информации о взаимных связях между признаками. Т.е. отсутствует использование наиболее подходящих для этого многомерных методов. Об этой ситуации мы уже писали в разделе «**ДОКАЗАТЕЛЬНАЯ МЕДИЦИНА И СТАТИСТИКА**» [41] нашего обзора «Доказательная или сомнительная? Медицинская наука Кузбасса: статистические



Прочитай и перешли своим коллегам…

аспекты». В частности, для иллюстрации ошибочного представления о достаточности такого подхода при сравнении групп, в обзоре приведён
представленный слева график. Для данных групп 1 и 2 средние значения совокупностей, откуда извлечены эти выборки, статистически значимо не различаются. Однако при таком «плоском» сравнении игнорируется информация, заложенная в корреляционных связях всех анализируемых признаков. Поэтому авторские выводы о сопоставимости традиционного биохимического метода и метода малоуглового рентгеновского рассеяния у больных ИБС, не очень надёжны и доказательны. Поскольку при аргументации такого вывода игнорируются взаимосвязи между показателями липидного спектра, а также некорректно используется t-критерий Стьюдента. Применительно к данной задаче и анализируемому массиву данных более разумно было бы использование таких многомерных методов статистического анализа, как непараметрический дискриминантный анализ и логистическая регрессия [44].

5) ПКК № 1, 2008. С. 18-21. Применение коронарных стентов с лекарственным покрытием в лечении больных с рецидивом стенокардии после коронарного шунтирования.

Авторы статьи пишут: «Достоверность различий анализировали с помощью t-критерия Стьюдента для выборок с **параметрическим распределением** в доверительном интервале более 95%». При этом авторы не уточняют, что такое «**параметрическое распределение**». Между тем из учебников теории вероятности и математической статистики известно, что **ВСЕ РАСПРЕДЕЛЕНИЯ имеют свои параметры**. Например, такие параметры как среднее, дисперсия, и т.д. Таким образом, авторы утверждают, что они применяли t-критерий Стьюдента **ДЛЯ ВСЕХ РАСПРЕДЕЛЕНИЙ**. Что является грубейшей ошибкой, поскольку данный критерий можно корректно применять лишь в весьма ограниченном числе случаев. В частности, лишь тогда, когда одновременно выполняются 2 ограничительных условия: нормальность распределения в обеих группах сравнения, а также равенство двух генеральных дисперсий в группах сравнения. Некорректно и использование выражения «достоверность различий» [24] применительно к проверке статистических гипотез с использованием выборок. Непонятен также и смысл упоминания о «доверительном интервале более 95%». Поскольку авторы проводили сравнение между собой 3-х групп, тогда как t-критерий Стьюдента позволяет сравнивать лишь 2 группы, то можно также говорить о некорректности подобных сравнений Т.к. авторы не учитывали проблему множественных сравнений.

6) ПКК № 2, 2010. С. 18-21. Применение аминодифосфоната для профилактики кальцификации эпоксиобработанных биопротезов.

> Полученные результаты обрабатывали с помощью программы «Statistica 6.0». Все данные представлены как средние значения (M) и стандартная ошибка среднего (±m). Статистическую значимость различий оценивали с помощью t-критерия Стьюдента. Достоверными считали различия при уровне значимости $p<0,05$.

Авторы не проверили ограничения на возможность применения t-критерия Стьюдента, и уже в силу этого их выводы можно считать сомнительными. Помимо этого авторы используют некорректный оборот «Достоверными считали различия при уровне значимости $p<0,05$» (см. фрагмент текста слева). Уже сам факт ненулевого значения критического уровня значимости в 5% говорит о том,

что полученные результаты в принципе не могут быть достоверными [24] (см. пояснения этой некорректности выше). Авторы проводят сравнение группы «Контроль» и 3-х временных точек: 60 суток, 120 суток и 180 суток. Наличие 4-х групп сравнение требует обязательного учёта проблемы множественных сравнений. Поскольку для 4 групп имеется 6 пар сравнения, то критический уровень значимости должен быть понижен в 6 раз. И в этом случае он будет менее 1%. Однако авторы не подозревают о данной проблеме и не изменили критический уровень значимости. К положительным моментам данной публикации следует отнести то, что авторы в разделе «Материал и методы» поясняют используемые ими выражения вида $M \pm m$, а также используют в тексте выражение о статистически значимых отличиях.

7) ПКК № 1, 2011. С. 16-20. Непосредственные результаты протезирования аортального клапана каркасными биологическими протезами «БиоЛАБ КА/ПТ».

В разделе «Материал и методы» авторы не сообщают, какие конкретно статистические методы были ими использованы. Между тем, в разделе «Результаты и обсуждение» они приводят результаты использования статистических методов. Так, в табл. 1 приводятся результаты сравнения средних четырёх показателей. При этом приведены значения достигнутого уровня статистической значимости для использованных статистических критериев. Однако уровни значимости авторы ошибочно называют уровнем достоверности. Выше мы уже обсуждали смысл этой некорректности [24]. Так, для показателя ИММЛЖ эта величина равна 0,00008. В тексте статьи часто приводятся значения средних величин, с использованием выражений вида 69,1 ± 4,5. Однако при этом авторы не сообщают, какая конкретно величина, ошибка среднего, стандартное отклонение, или же полуширина доверительного интервала размещена после знака ± . В статье сообщается, что «Для выявления факторов, оказывающих наибольшее влияние на формирование градиента давления на биопротезе «БиоЛАБ КА/ПТ», провели нелинейный многомерный статистический анализ данных». Но при этом в статье абсолютно ничего не говорится о том, какой же конкретно метод статистического анализа был использован авторами. Судя по приведённому далее выражению «ПТГД = 27,7±3,09 ИМТ^2 – 0,036 (№ протеза)^2» можно предположить, что ими был использован метод множественной регрессии с предикторами второго порядка.

В тексте статьи сообщается: «При проведении анализа выявили, что существует сильная ($p<0,05$) зависимость между величиной ПТГД и ИМТ, № протеза и площадью отверстия протеза». Данное утверждение также является ошибочным, поскольку сила, интенсивность связи в множественной регрессии, определяется не величиной достигнутого уровня статистической значимости, а значением такого показателя, как коэффициент множественной корреляции R и коэффициент детерминации R^2. Чем ближе эти величины к 1, тем сильнее связь. Например, для двух разных уравнений множественной регрессии были получены одинаковые значения достигнутого уровня значимости для F-критерия Фишера, равные 0,01. Однако для первого уравнения регрессии коэффициент детерминации был равен 0,81, а для второго он был равен 0,49. В этом случае можно говорить о том, что в первом уравнении 81% вариабельности зависимой переменной определяется вариабельностью вошедших в уравнение предикторов. И только 19% изменчивости зависимой переменной зависят от иных величин, не вошедших в уравнение. Тогда как во втором случае изменчивость предикторов определяет лишь 49% вариабельности величины

зависимой переменной. А остальные 51% зависят от иных, не вошедших в уравнение признаков. В этом случае можно говорить о наличии более сильной связи для первого уравнения, и менее сильной – для второго. Более того, величины R и R^2 существенно корректируются (уменьшаются) с учётом объёма наблюдений, по которым оцениваются уравнение множественной регрессии. А эти объёмы для каждого уравнения могут значительно отличаться между собой.

Вызывает вопрос и форма приводимого в статье уравнения: «ПТГД = 27,7 ± 3,09 $ИМТ^2$ – 0,036 (№ протеза)^2» . Каков смысл числа 3,09 в выражении 27,7 ± 3,09? И почему второй коэффициент 0,036 не имеет такую же форму? Т.е. нет знака ± после 0,036 и после него нет второго числа. Скорее всего, это опечатка, пропущенная ответственными за данный выпуск журнала редактором Т. Ф. Чалковой и корректором Н. Ф. Подопригорой. Т.е. число 27,7 – это свободный член уравнения (Intercept), а вместо знака ± должен быть один знак, а именно знак +, поскольку в статье авторы пишут о прямом (читай: положительном) влиянии индекса массы тела. И в этом случае полученное авторами уравнение множественной регрессии, видимо, имеет следующий вид: ПТГД = 27,7 + 3,09* $ИМТ^2$ – 0,036*(№ протеза)^2.

При использовании метода множественной регрессии оценивается довольно много величин. В частности, это размерные регрессионные коэффициенты, безразмерные регрессионные коэффициенты, стандартные ошибки этих коэффициентов, критерии проверки статистической значимости этих коэффициентов, достигнутые уровни значимости для каждого коэффициента, включая так называемый свободный член (Intercept), коэффициент множественной корреляции (интенсивность, сила связи), коэффициент детерминации (интенсивность, сила связи), F-критерий Фишера и 2 степени свободы для данного статистического критерия. Тогда как применительно к уравнению регрессии авторы статьи лишь однажды упоминают уровень значимости: «… выявили, что существует сильная ($p<0,05$) зависимость …». А какова «сила» этой зависимости, авторы не сообщают. Непонятно также и то, что означает признак «№ протеза». Если между разными номерами протезов имеются различия в их параметрах (линейные размеры, размеры площади и объёма, и т.д.), то вполне разумно было бы ввести набор этих характеристик в список потенциальных регрессоров (предикторов). Т.е. в состав признаков, которые могли войти в уравнение множественной регрессии для объяснения величины зависимого показателя ПТГД. Однако в тексте статьи об этом ничего не сказано.

Другой аспект использования многомерных статистических методов заключается в том, что в этом случае априорно не может быть лишь одного единственного решения. Так, для случая множественной регрессии возможна оценка не одного, а целого набора уравнений. Тем более, в случае нелинейного уравнения. Во-первых, имеется огромное количество нелинейных преобразований, из которых тем или иных способом отбираются уравнения, дающие лучшие результаты. Между тем, авторы никак не аргументируют своего выбора именно уравнения второго порядка, а не иного нелинейного преобразования. Во вторых, оценка таких уравнений может быть произведена с помощью различных алгоритмов. Так, помимо принудительного включения предикторов в уравнение достаточно плодотворно используются так называемые пошаговые алгоритмы оценки уравнений. В третьих, в ряде случаев полученные уравнения не всегда удовлетворяют тем ограничениям, которые накладываются на этот метод. И в этом случае

необходимо оценивать устойчивость параметров уравнений, а также исследовать анализируемые данные на предмет наличия в них так называемых аномальных значений (выбросов). Обо всех этих аспектах авторы статьи ничего не сообщают, видимо, по причине незнания этих деталей данного метода. В итоге полученное авторами уравнение множественной регрессии следует оценивать как недостаточно обоснованное, а значит, его использование в практике вполне может привести к негативным последствиям.

8) ПКК, № 1, 2010, с. 3-6. Морфология компенсаторно-приспособительной перестройки миокарда у кардиохирургических больных с транспозицией магистральных сосудов.

Авторы пишут: «Статистическая обработка результатов проводилась с помощью программы Excel. Достоверность различий средних величин и корреляционных взаимоотношений проводили с помощью t-критерия Стьюдента. Достоверными считали различия при p<0,05». В работе допущена типичная ошибка – использование критерия Стьюдента без проверки ограничений его использования. Напомним, что в пакете Excel невозможно проверить в полном объёме эти ограничения. Кроме того, два ограничения на корректное использование данного критерия одновременно выполняются лишь в нескольких процентах случаев, что уже делает авторские выводы сомнительными. Далее, авторы используют в своей работе 3 группы сравнения: «До 1 мес», «1-6 мес» и «6-12 мес». В этом случае необходимо учитывать проблему множественных сравнений, и по этой причине уменьшать используемый в статье критический уровень значимости p=0,05. В статье авторы используют некорректные обороты со словом «достоверно» [24] (см. пояснения выше). В работе используются выражения для вычисленных средних величин вида «45,2±2,2», однако при этом авторы не уточняют, какие величины размещаются после знака ± .

9) ПКК, № 1, 2010, с. 7-12. Результаты послерадикальной коррекции тетрады Фалло с позиций хирургической тактики при реконструкции выходного отдела правого желудочка.

В этой статье авторы используют критерий Стьюдента без проверки ограничений на его применение, а также некорректные обороты со словом «достоверно» [24]. Положительным элементом можно считать следующий фрагмент текста: «Критическая величина уровня значимости (p) была принята равной 0,05. Нулевая гипотеза отвергалась, когда значение p было менее 0,05». В работе используются выражения для вычисленных средних величин вида «2,43±0,57», однако при этом авторы не уточняют, какие величины размещаются после знака « ± ». В таблице 2 авторы приводят результаты вычисления одностороннего точного критерия Фишера для двух таблиц 2x2 со следующими частотами:

Таблица 2

Типы и частота встречаемости диастолических дисфункций ПЖ после РКТФ

Группы	Замедление релаксации		Рестриктивный тип		p
	абс.	%	абс.	%	
I	22	23,5	15	16	0,45
II	30	32	26	28	
IIA	17	30	18	32	0,24
IIB	13	23	8	14	

Однако расчёт одностороннего точного критерия Фишера для таблицы с частотами

22 15

30 26

даёт иной результат: p=0,365, вместо приведённого результаты p=0,45.

Прочитай и перешли своим коллегам…

А для таблицы
17 18
13 8
получаем результат p=0,245. Разумеется, в обоих случаях полученные вероятности гораздо больше критического уровня p=0,05 , однако большая разница в результатах для первой таблицы вызывает сомнения в корректности других результатов. К недостаткам данного исследования следует отнести также и ограниченность использованных методов статистического анализа. В частности, анализировались лишь парные связи, при игнорировании методов многомерного статистического анализа, которые позволяют получать более ценную и полезную для целей исследования информацию. Учитывая наличие 2 и 4 групп сравнения, целесообразно было использовать в данном исследовании, как минимум, такие методы анализа, как факторный анализ [см. Факторный анализ: основные положения и ошибки применения http://www.biometrica.tomsk.ru/factor.htm], логистическую регрессию [44], непараметрический дискриминантный анализ.

10) ПКК, № 1, 2010, с. 13-20. Особенности ультраструктуры эндотелия коронарных микрососудов при хирургической коррекции врожденного порока сердца у детей первого года жизни и в возрасте одного–трёх лет.

Положительный момент данной публикации: конкретизация структуры выражения вида 8,8 ± 0,7. Отрицательные моменты: использование критерия Стьюдента без проверки ограничений его использования, а также без учёта проблемы множественных сравнений, и использование оборотов со словом «достоверно» [24]. К недостаткам данного исследования следует отнести также и ограниченность использованных методов статистического анализа. В частности, анализировались лишь парные связи, при игнорировании методов многомерного статистического анализа, которые позволяют получать более ценную и полезную для целей исследования информацию. Учитывая наличие 2 групп сравнения, целесообразно было использовать в данном исследовании, как минимум, такие методы анализа, как логистическую регрессию [44], ROC-анализ [69-70], непараметрический дискриминантный анализ.

11) ПКК, № 1, 2010, с. 21-25. Протезирование аортального клапана: прогноз результатов.

Авторы сообщают, что «Цель исследования – оценка влияния исходных анатомо-функциональных и гемодинамических показателей на прогнозирование ближайших результатов у пациентов после протезирования АК». Далее в разделе «Материал и методы» авторы приводят следующую информацию. «Для оценки влияния исходных анатомо-функциональных показателей на результаты операции протезирования АК проведено клинико-функциональное исследование 394 пациентов, которым было выполнено изолированное протезирование АК за период 2001–2007 гг.. … Анализировались следующие исходные показатели:

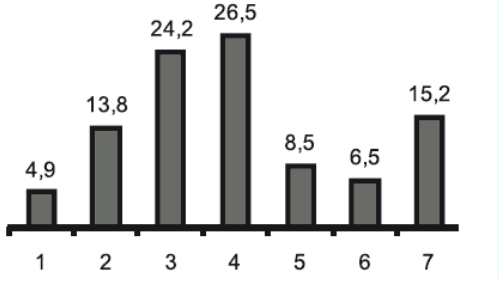
1. пол,
2. возраст (годы),
3. вес (кг),
4. рост (см),
5. площадь поверхности тела (кг/м2),
6. индекс массы тела (Кетле, отн. ед.),
7. стадия хронической сердечной недостаточности (ХСН),
8. степень функционального класса по NYHA,

Прочитай и перешли своим коллегам…

9. степень аортального стеноза,
10. степень активности инфекционного эндокардита,
11. систолическое (САД),
12. диастолическое (ДАД) и артериальное давление (мм рт. ст.),
13. ЧСС (уд/мин),
14. КДР и
15. КСР ЛЖ (см),
16. КДО и
17. КСО ЛЖ (мл),
18. УО ЛЖ (мл);
19. ФВ ЛЖ (%),
20. фракция укорочения ЛЖ (%),
21. толщина задней стенки ЛЖ в диастолу (ТЗСЛЖд, см),
22. толщина межжелудочковой перегородки в диастолу (ТМЖПд, см),
23. масса миокарда левого желудочка (ММлж, г),
24. степень кальциноза АК (АКСа),
25. степень регургитации на АК (АКрег),
26. митральном клапане (МКрег),
27. трикуспидальном клапане (ТКрег) в баллах от 0 до 4 с дискретностью 0,5;
28. диаметр корня и восходящего отдела аорты (см);
29. кардиоторакальный индекс (КТИ,%);
30. пиковый и средний систолический градиент давления на АК (мм рт. ст.),
31. пиковый градиент давления на МК;
32. площадь митрального отверстия (МО, см2);
33. величина пика Е и А на МК (Е МК, А МК) и
34. отношение Е/А МК (отн. ед.).

Кроме того, рассчитывался ряд дополнительных показателей [3]. Для всех этих показателей был проведён анализ корреляционных взаимоотношений, оценивая существующие связи с результатом операции, с учётом их достоверности, направленности и силы. … Учитывая многообразие причинно-следственных связей и параметров, характеризующих состояние больных с пороками АК и необходимость создания модели, на основе которой возможно прогнозировать течение послеоперационного периода и **вероятность благоприятного исхода операции**, мы применили метод многомерного статистического анализа. Модель прогноза рассчитана на базе **регрессионного анализа**, задачей которого являлось прогнозирование значений результирующей переменной Y по известным значениям антропометрических параметров, показателей ЭхоКГ и ряда дополнительных характеристик, связанных со спецификой проведения операций на дооперационном этапе. Расчёты проводились методами математической статистики».

Выше приведены 34 показателя, из которых 8 показателей имеют балльную шкалу. Кроме того, как сообщают авторы, «рассчитывался ряд дополнительных показателей [3]». В статье же не приводится ни одного уравнения регрессии, хотя об использовании этого метода анализа авторы заявили в разделе «Материал и методы».



1 – НК; *2* – антропометрические показатели; *3* – показатели ЦГД; *4* – параметры сердца; *5* – показатели миокарда; *6* – морфология клапанов; *7* – показатели клапанов.

Рис. 1. Доля влияния факторов на прогноз, %.

«В ходе исследования было выявлено, что для пациентов с изолированным протезированием АК большее влияние на прогноз операции играли факторы параметров сердца, центральной гемодинамики, показатели клапанов сердца, антропометрические данные и показатели миокарда (рис. 1)». Однако в статье не объясняется, каким конкретно методом

статистического анализа были получены эти оценки «доли влияния факторов на прогноз, %». Если предположить, что авторы использовали для этой цели заявленный ими регрессионный анализ, то в этом случае необходимо уточнить, какая величина принималась в качестве результирующей переменной Y. В подписи к рис. 1 авторы упоминают такой показатель, как прогноз. Однако в списке 34 перечисленных в разделе «Материал и методы» признаков такой показатель отсутствует. Иными словами, неясно, что за величину авторы статьи подразумевают под прогнозом. Если это величина количественная, то в этом случае разумно привести полученное уравнение множественной регрессии, хотя бы с минимумом параметров: коэффициенты регрессии (размерные и безразмерные), достигнутые уровни статистической значимости для коэффициентов, показатели адекватности уравнения в целом (коэффициент множественной корреляции, коэффициент детерминации, F-критерий Фишера и достигнутый уровень значимости для него).

В статье приведено несколько десятков коэффициентов парной корреляции. «Анализ показателей Ф3 демонстрирует умеренную корреляционную связь гемодинамических показателей с прогнозом операции (r=0,424). При этом влияние показателей на прогноз операции было больше в группе пациентов с АоН+СоАоН (r=0,232), чем в подгруппе с АоС+СоАоС (r=0,124)» «В ходе корреляционного анализа связи факторов с прогнозом операции были установлены следующие закономерности. Выявлена умеренная сила корреляции показателей Ф1 (r=0,683) с прогнозом операции, …». Напомним, что «показатели Ф1», это «фактор нарушения кровообращения (Ф1)». В статье авторы сообщают, что «Основные исходные показатели по близости характеризуемых ими переменных были объединены в группы факторов (Ф) …». Таких наборов было создано авторами 7. Из этого следует, что Ф1 есть не уникальный количественный или ранговый признак, а некий набор признаков. Однако авторы при этом не уточняют, какие конкретно признаки они включили в этот набор. Неясно и то, что за величина подразумевается под признаком «прогноз операции». Если это некая количественная величина, то в этом случае должен применяться один набор алгоритмов для оценки уравнения регрессии. Если же под прогнозом подразумевается качественный признак с двумя или более градациями, то в этом случае необходимо использовать логистическую регрессию [44]. В тексте статьи неоднократно встречаются аналогичные фразы, содержащие упоминание о связи прогноза операции и того или иного фактора. Однако нигде не конкретизируется, что такое «прогноз операции», и с помощью какого статистического метода авторы получили приводимые ими оценки корреляционных связей. Иными словами, для всех приводимых в статье коэффициентов корреляции, которых в статье около 50, не сообщается, что это за коэффициент корреляции. Напомним, что основных коэффициентов корреляции, отличающихся как по алгоритму их оценки, так и по комбинациям анализируемых признаков, в математической статистике известно порядка десятка. Это коэффициенты корреляции Пирсона, Спирмена, Кендэла, Фехнера, коэффициент множественной корреляции, двухстрочечной корреляции, бисериальной корреляции и т.д.

«Выявление показателей, исходные значения которых могут определять процент благоприятного исхода операции, прогнозировать возможные осложнения, равно как и оценка тяжести состояния в предоперационном периоде у пациентов, идущих на протезирование АК, является одним из актуальных направлений современной кардиохирургии». Поскольку в приведённом выше

фрагменте статьи сообщается об определении процента благоприятного исхода, то весьма вероятно, что использована логистическая регрессия. Но тогда результаты такого анализа должны содержать помимо самого уравнения с коэффициентами и уровнями значимости, и другие показатели, в частности, показатели конкордации, ROC-кривые [44, 69-70].

Авторы сообщают, что «В нашем исследовании корреляционному анализу были подвержены 68 исходных анатомофункциональных показателя». Если же принять во внимание, что часть анализируемых признаков имела балльную шкалу, то очевидно, что в этом случае не всегда возможно использовать классические алгоритмы множественного регрессионного анализа. Более приемлем в этом случае более сложные методы, такие как метод логистической регрессии, лог-линейный анализ, анализ соответствий (correspondence analysis) , анализ многомерных таблиц сопряжённости, и т.д. Кроме того, с учётом того, что несколько десятков показателей являются количественными, для достижения поставленных целей исследования было бы разумно использовать и такие методы статистического анализа, как факторный анализ и анализ канонических корреляций. В итоге можно сделать заключение, что из ценнейшего массива, содержащего значения нескольких десятков показателей по 394 пациентам, извлечена лишь самая мизерная толика информации. Причём корректность результатов этого анализа вызывает большие сомнения.

12) ПКК, № 1, 2010, с. 35-37. Изменение систолической и диастолической функции левого желудочка у пациентов с ишемической болезнью сердца в раннем периоде после коронарного стентирования по данным допплер –эхокардиографии.

К недостаткам публикации следует отнести использование оборотов со словом «достоверно» [24]. Кроме того, авторы не уточнили, для проверки на соответствие какому конкретно виду распределения был использованы критерий Колмогорова-Смирнова. Можно предположить, что в этом случае подразумевалось нормальное распределение, однако в этом случае использование данного критерия имеет свою специфику, заключающуюся в том, что для вычисления величины критерия используются выборочные параметры распределения, и для корректировки критерия необходимо использовать соответствующие поправки. Кроме того, не сообщается и о проверке второго ограничительного условия для использования t-критерия Стьюдента. К положительным моментам публикации следует отнести пояснение структуры выражения для средних величин вида 55,3±9, а также факт упоминания критического уровня статистической значимости, и использованных статистических критериев сравнения центральных мер.

13) ПКК, № 1, 2010, с. 41-43. Хирургический стресс и белки теплового шока HSP70 при операциях аортокоронарного шунтирования.

В тексте статьи авторы используют выражения вида 63,8±3,5, однако не сообщают, какие конкретно величины соединены знаком ± . Если относительно первой величины можно быть уверенным, что это есть среднее, то относительно второй величины, стоящей после знака ± такой уверенности нет. Поскольку это может быть как ошибка среднего, так и стандартное отклонение,

а также полуширина доверительного интервала. В редакционных же требованиях нет никаких указаний о том, что за величины могут представляться в такой форме.

Таблица 1

Содержание кортизола и БТШ HSP70 у больных ИБС на этапах реваскуляризации миокарда в условиях искусственного кровообращения

Этап	Кортизол, нмоль/л	HSP70, нг/мл
Исходный до операции	356±31,9	0,08±0,02
Окончание операции	990±68,9***	0,41±0,06***
Через 6 ч после окончания	1921±187***	0,31±0,09*
1-е сутки после операции	699±73,3***	0,096±0,01
2-е сутки после операции	695±71,4***	0,08±0,01

** p<0,05; *** p<0,001 – различия достоверны по сравнению с исходным этапом*

Таблица 2

Содержание тропонина I и активность миокардиальной фракции креатинкиназы у больных ИБС на этапах реваскуляризации миокарда в условиях искусственного кровообращения

Этап	Тропонин I нг/мл	Активность КК-МВ, Ед/л
Исходный до операции	0,53±0,20	13,8±1,89
Окончание операции	1,43±0,32*	36,8±3,3***
Через 6 ч после окончания	2,76±0,55***	31,5±3,35***
1-е сутки после операции	2,24±0,39***	37,1±0,01***
2-е сутки после операции	1,39±0,33*	39,2±8,2**

** p<0,05; ** p<0,01; *** p<0,001 – различия достоверны по сравнению с исходным этапом*

Учитывая тот факт, что в статье использованы выражения вида $p < 0,05$ и $p < 0,01$ (см. табл. 1 и 2, а также текст статьи), а также некорректные обороты вида «различия достоверны» [24], можно предположить, что авторы статьи использовали некие статистические критерии. Именно предположить, а не утверждать. Поскольку в разделе «Материал и методы» авторы не сочли необходимым указать, какие конкретно статистические методы и критерии были ими использованы.

14) ПКК, № 1, 2010, с. 44-48. Ретроградная перфузия головного мозга как компонент противоишемической защиты мозга при реконструктивных операциях на дуге аорты.

В отличие от предыдущей статьи, в данной статье в разделе «Материал и методы» авторы перечисляют использованные ими методы и указывают критический уровень статистической значимости, а также сообщают, какие конкретно величины соединены знаком « ± »: «Статистический анализ проведён с помощью программ Excel, Биостат и Statistica 6. Сравнение двух групп проводили с применением t-критерия Стьюдента. Качественные признаки оценивались с использованием критерия χ^2. Результаты представлены как среднее и стандартная ошибка среднего (M ± m). Статистически значимыми считались различия при $p<0,05$». Однако говоря об использовании критерия Стьюдента, авторы не указывают, каким образом проверялись ограничения на его использование. Из чего можно сделать вывод о том, что данные ограничения им неизвестны. Кроме того, учитывая, что число групп сравнения более двух, нужно было учитывать проблему множественных сравнений, для чего необходимо уменьшать значение критического уровня значимости. В статье же об учёте этой проблемы авторы не сообщают. В работе используются некорректные выражения вида «***при уровне достоверности * p<0,05; ** p<0,01***» [24]. То, что авторы статьи называют неверно уровнем достоверности, в действительности есть величина достигнутого уровня статистической значимости «p» [65]. Не корректно также выражение «Качественные признаки оценивались с использованием критерия $\chi 2$». Например, если таким признаком предположить пол пациента, то очевидно, что пол пациента не оценивается с помощью критерия χ^2 . В данном выражении авторы подразумевали, что критерий Хи-квадрат использовался ими для анализа статистических связей между качественными признаками. В

разделе «Обсуждение» авторы приводят подобные результаты анализа: «В нашем исследовании включение в схему защиты мозга РПГМ позволило снизить летальность (РПГМ – 3,4%, ГГОК – 25%; χ^2, p=0,002) и распространённость послеоперационной неврологической дисфункции (РПГМ – 7,9%, ГГОК – 37,5%; χ^2, p<0,05)».

Отметим, что данную форму приведения результатов анализа статистической связи между качественными признаками нельзя признать полноценной и информативной для читателя. В качестве образца более информативного представления результатов сравнения можно рассмотреть представленную ниже табл. 2.

Таблица 2

Особенности течения послеоперационного периода в исследованных группах

Группы	Восстановление сознания, ч	Длительность ИВЛ, ч	Время в ОРИТ, сут.	Послеоперационный койко-день
РПГМ (n=89)	4,0±0,18	18,4±1,3	4,1±0,5	20,4±1,0
Подгруппа 1 (n=55)	4,0±0,19	17,1±1,5	3,9±0,6	19,4±1,2
Подгруппа 2 (n=34)	4,0±0,3	20,5±2,4	4,5±0,7	22,0±1,9
ГГОК (n=24)	5,2±0,5*	21,0±3,8	7,0±1,3*	25,7±2,9*

** p<0,05 отличие по сравнению с группой РПГМ*

В этой таблице читатель достаточно наглядно видит изменение средних величин наиболее важных показателей. Тогда как выражения вида «РПГМ – 3,4%, ГГОК – 25%; χ^2, p=0,002» не обладают такой наглядностью. Поскольку авторы статьи не приводят ни значения самого критерия Хи-квадрат, ни числа степеней свободы для этого критерия, то трудно решить, при анализе таблицы сопряжённости какого вида был получен этот уровень статистической значимости. Ведь достигнутый уровень значимости p=0,002 может быть получен как для разных значений самого критерия Хи-квадрат с разными степенями свободы df. То есть, иными словами, для разных таблиц сопряжённости. Так достигнутый уровень значимости p=0,002 получаем для значения критерия χ^2=9,54936, и при числе степеней свободы равном 1, т.е. для таблицы сопряжённости 2x2. И этот же самый уровень значимости p=0,002 получаем для таблицы сопряжённости 2x3, при числе степеней свободы равном 2, и величина критерия χ^2 при этом будет равна 12,429216. Однако как в случае таблицы 2x2 (4 клетки), так и в случае таблицы 2x3 (6 клеток), статистически значимая взаимосвязь между двумя качественными признаками не будет равномерно распределена («размазана») по всем 4 или 6 клеткам (см. подобный пример в нашей статье «ДОКАЗАТЕЛЬНАЯ МЕДИЦИНА И СТАТИСТИКА [41]). Тщательный анализ позволяет установить, какие конкретно клетки (строки, столбцы) в большей степени ответственны за наличие значимой связи. Иными словами, для вдумчивого, интеллектуального читателя, таким же интеллектуальным авторам целесообразнее не просто приводить значение уровня значимости p=0,002, а приводить и всю таблицу сопряжённости. И, кроме того, желательно приводить и детальный анализ такой таблицы. Естественно, что для этого необходимо располагать специальными знаниями по математической статистике, что, в свою очередь, требует участие в анализе подобных данных профессионального биостатистика.

15) ПКК, № 1, 2010, с. 56-59. Факторы риска госпитальной и отдалённой летальности хирургического лечения больных ишемической болезнью сердца с выраженной дисфункцией левого желудочка.

В разделе «Материал и методы авторы пишут: «Для проверки гипотезы применялся критерий χ^2». Однако содержание текста статьи говорит также и об использовании иных статистических критериев.

Таблица 1

Результаты стрессЭхоКГ с малыми дозами добутамина

Группы		КДО	КСО	ФВ	p
Полож. ответ (n=43)	Исходно	233±43	158±38	32±7	>0,05
	Стресс-ЭхоКГ	217±11	131±0,7	40±3	
Отриц. ответ (n=24)	Исх.	251±25	174±16	31±5	>0,05
	Стресс ЭхоКГ	243±0,8	172±0,4	31±2	

Таблица 2

Динамика ФВ ЛЖ после АКШ в зависимости от результатов стресс-ЭхоКГ

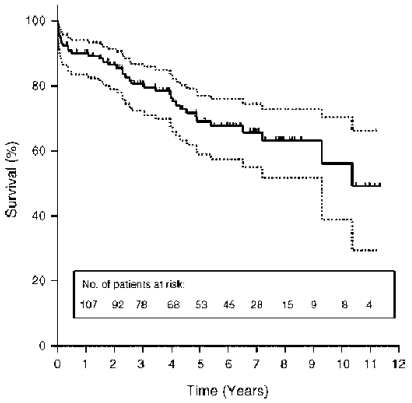
Группы		КДО	КСО	ФВ	p
Полож. ответ	Исходно	233±43	158±38	32±7	<0,05
	Через 4 мес.	228±11	143±8	37±4	
Отриц. ответ	Исходно	251±25	174±16	31±5	>0,05
	Стресс-ЭхоКГ	247±6	171±4	30±4	

Слева приведены табл. 1 и 2, в которых, видимо, приводятся средние значения 4-х групп сравнения, а также уровни значимости. Однако авторы не сообщают, каким методом и критерием они пользовались для сравнения 4-х групп. Не сообщают авторы и о том, какие выборочные параметры размещены после знака «±» в выражениях вида 233 ± 43. В тексте используются некорректные обороты со словом «достоверно» [24]. «Статистический анализ логистической регрессии выявил, что объемные показатели и фракция выброса ЛЖ по данным ЭхоКГ не влияли на отдаленную летальность». Таким образом, авторы сообщают об использовании ими метода логистической регрессии. Напомним, что существуют разные модели и алгоритмы логистической регрессии, результаты которых существенно отличаются между собой [44]. Однако в статье не уточняются использованные авторами алгоритмы логистической регрессии.

Авторы дают ссылку на источник [11]: Sartipy U., Albage A., Lindblom D. // Eur. J. Cardiothorac. Surg. 2006. V. 30. P. 762–771. Если же обратиться к этой статье, то можно обнаружить, что в ней авторы гораздо более подробно описывают полученные ими результаты. Например, в таблице 4 они приводят следующую информацию

Table 4
Bivariate analysis of preoperative and surgical risk factors for early mortality in all patients (n = 136)

Variables	Early mortality Mean ± SD or percentage (n)		p
	Yes (n = 10)	No (n = 126)	
Preoperative variables			
Age (years)	67.2 ± 9.2	63.7 ± 9.7	0.24
Female	30% (3)	23% (29)	0.70
Ejection fraction	23.5 ± 8.2	26.6 ± 9.8	0.38
NYHA class III–IV	90% (9)	68% (86)	0.28
MR grade III–IV	30% (3)	12% (15)	0.13
Left main stenosis	10% (1)	2.4% (3)	0.27
Inducible VT	20% (2)	40% (50)	0.32
Spontaneous VT	30% (3)	24% (30)	0.70
Tripple vessel disease	70% (7)	49% (62)	0.33
Anterior LVA	80% (8)	92% (116)	0.22
Posterior LVA	10% (1)	4.8% (6)	0.42
IDCMP	10% (1)	4.0% (5)	0.37
Previous cardiac surgery	30% (3)	3.2% (4)	0.009
BMI (kg/m^2)	24.4 ± 4.0	26.1 ± 3.8	0.24
Diabetes mellitus [a]	40% (4)	17% (21)	0.09
Impaired renal function [b]	40% (4)	18% (22)	0.09
Euroscore	9.2 ± 2.6	7.1 ± 2.8	0.02



(См. таблицу 4 слева). Как видим, в таблице 4 используется выражение «Mean ± SD», т.е. среднее ± стандартное отклонение, а также значения достигнутых уровней значимости.

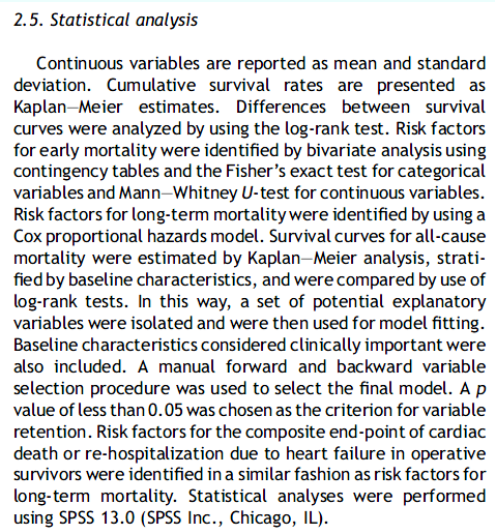
Table 5
Risk factors for late mortality (n = 136) and composite endpoint of first postoperative hospital re-admission due to heart failure or cardiac death in operative survivors (n = 126) after surgical ventricular restoration by Cox proportional hazards analysis

Outcome and variables	Level	HR	95% CI	p
Late mortality (n = 136, number of events = 40)				
Age	One year	1.05	1.01–1.1	0.02
Diabetes mellitus[a] (n = 25)	Yes/no	2.75	1.4–5.5	0.004
Mitral regurgitation grade III–IV (n = 18)	Yes/no	2.83	1.3–6.1	0.008
Hospital re-admission/death (n = 126, number of events = 49)				
Age	One year	1.06	1.02–1.1	0.002
Mitral regurgitation grade				
0 (n = 57)	Yes/no	1 (reference)		
I (n = 24)	Yes/no	1.49	0.6–3.6	0.38
II (n = 30)	Yes/no	1.97	0.9–4.1	0.07
III (n = 11)	Yes/no	2.65	1.0–6.8	0.04
IV (n = 4)	Yes/no	4.59	1.2–17	0.02

CI: confidence interval, HR: hazard ratio.
[a] Treated with insulin or oral anti-diabetics.

А вот какие подробные результаты приводят авторы этой статьи по оценке факторов риска методом Кокса (см. слева):

Прочитай и перешли своим коллегам…



2.5. Statistical analysis

Continuous variables are reported as mean and standard deviation. Cumulative survival rates are presented as Kaplan–Meier estimates. Differences between survival curves were analyzed by using the log-rank test. Risk factors for early mortality were identified by bivariate analysis using contingency tables and the Fisher's exact test for categorical variables and Mann–Whitney *U*-test for continuous variables. Risk factors for long-term mortality were identified by using a Cox proportional hazards model. Survival curves for all-cause mortality were estimated by Kaplan–Meier analysis, stratified by baseline characteristics, and were compared by use of log-rank tests. In this way, a set of potential explanatory variables were isolated and were then used for model fitting. Baseline characteristics considered clinically important were also included. A manual forward and backward variable selection procedure was used to select the final model. A p value of less than 0.05 was chosen as the criterion for variable retention. Risk factors for the composite end-point of cardiac death or re-hospitalization due to heart failure in operative survivors were identified in a similar fashion as risk factors for long-term mortality. Statistical analyses were performed using SPSS 13.0 (SPSS Inc., Chicago, IL).

Слева также приведён фрагмент этой статьи с описанием раздела «Статистический анализ».
Данные примеры приведены нами для того, чтобы показать, что статья, на которую делают ссылку авторы, могла быть образцом для подражания в части более полного описания результатов использования статистических методов.

16) ПКК, № 2, 2010, с. 4-7. Протезирование митрального клапана с сохранением подклапанных структур у пациентов детского возраста.

В работе используются выражения для вычисленных средних величин вида «11,4±4,8», однако при этом авторы не уточняют, какие величины соединены знаком « ± ».
В статье используются выражения вида «достоверные различия» [24], что свидетельствует о незнании авторами статьи основных понятий и терминологии статистики [65]. Авторы не сообщают, с помощью каких конкретно статистических методов и критериев они проверяли гипотезы о различии средних 9 признаков для трёх групп сравнения.

17) ПКК, № 2, 2010, с. 8-12. Отсроченное ушивание стернотомной раны у новорожденных и детей раннего возраста после коррекции транспозиции магистральных сосудов.

В статье авторы пишут: «Статистический анализ проводился с использованием критерия Стьюдента, одностороннего **точечного критерия Фишера**, достоверный уровень значимости «р» был принят 0,05». Однако при этом не проводят проверку ограничений на использование критерия Стьюдента, что уже ставит под сомнение все полученные с его использованием результаты. Некорректно также и выражение «точечный критерий Фишера». Поскольку в статистике нет «точечного критерия Фишера». А есть «точный критерий Фишера». Предположим, что авторы допустили элементарную опечатку. Однако при этом ни рецензенты, ни редколлегия журнала, ни ответственные за выпуск данного номера журнала, также не оспорили некорректность этого выражения. Это наглядно свидетельствует как о качестве статистического анализа проводимого авторами, так и о качестве рецензирования, и работы с рукописями редколлегии ПКК. Некорректно также использование выражения «достоверный уровень значимости «р» был принят 0,05» и остальных выражений с использованием слова «достоверно» [24]. В работе используются выражения для вычисленных средних величин вида «3,6±0,9», однако при этом авторы не уточняют, какие величины соединены знаком « ± ».

18) ПКК, № 2, 2010, с. 18-21. Применение аминодифосфоната для профилактики кальцификации эпоксиобработанных биопротезов.

Прочитай и перешли своим коллегам…

В разделе «Материал и методы авторы пишут: «Полученные результаты обрабатывали с помощью программы «Statistica 6.0». Все данные представлены как средние значения (M) и стандартная ошибка среднего (±m). Статистическую значимость различий оценивали с помощью t-критерия Стьюдента. Достоверными считали различия при уровне значимости p<0,05». Положительным моментом данного описания можно считать описание выражения M ± m, а также использование оборота «статистическая значимость». Однако помимо данного оборота авторы используют некорректные выражения со словом «достоверность» [24]. Поскольку авторы проводили с помощью критерия Стьюдента сравнение более 2-х групп, то в этом случае необходимо учитывая проблему множественных сравнений уменьшить критический уровень значимости, чего авторы не сделали. Далее, используя критерий Стьюдента, авторы не проводят проверку ограничений на применение данного критерия. Эти нарушения не позволяют считать надёжными выводы данной статьи относительно наличия статистически значимых различий средних.

19) ПКК, № 2, 2010, с. 22-25. Распределение химических элементов в сосудистой системе больных ишемической болезнью сердца.

В разделе «Материал и методы» авторы пишут: «Статистическая обработка полученных результатов проводилась с помощью программы Microsoft Excel 2000. **Достоверность отличий** рассчитывали по t-критерию Стьюдента. Достоверными считали различия при p<0,05. Результаты представлены в виде среднее значение ± ошибка среднего (M ± m)». Положительным моментом данного описания можно считать описание выражения M ± m, а также использование оборота «статистическая значимость». Однако помимо данного оборота авторы используют некорректные выражения со словом «достоверность» [24]. Поскольку авторы проводили с помощью критерия Стьюдента сравнение более 2-х групп, то в этом случае необходимо учитывая проблему множественных сравнений уменьшить критический уровень значимости, чего авторы не сделали. Далее, используя критерий Стьюдента, авторы не проводят проверку ограничений на применение данного критерия. В частности, в пакете EXCEL нет процедуры проверки нормальности распределения. Эти нарушения делают выводы данной статьи относительно наличия статистически значимых различий средних сомнительными. Поскольку авторы изучали парные корреляционные взаимосвязи между 13 химическими элементами, и часть коэффициентов корреляции была статистически значима, то разумно в этом случае было использовать многомерные методы, такие, как факторный анализ, кластерный анализ, дискриминантный анализ, логистическую регрессию и т.д. Это позволило бы извлечь более ценную информацию, нежели результаты сравнения средних 3-х групп наблюдений, и парные коэффициенты корреляции. **Учитывая огромную важность цели данного исследования, предлагаем авторам данной статьи бесплатное проведение детального статистического анализа собранных ими данных, и дальнейшее сравнение полученных результатов анализа с опубликованными в данной статье.**

20) ПКК, № 2, 2010, с. 26-28. Влияние гепарина и протамина сульфата на агрегацию эритроцитов.

В разделе «Материал и методы» авторы пишут: «Данные исследования обработаны непараметрическими методами вариационной статистики с использованием программы «STATISTICA 6.0». Результаты представлены в виде M ± m. Различия считали достоверными при p < 0,05». Авторы не уточняют, какие конкретно величины связаны символом « ± ». Не

сообщается также и о том, какие конкретно непараметрические методы были использованы. К недостаткам статьи можно отнести также использование некорректных оборотов со словом «достоверно» [24].

21) ПКК, № 2, 2010, с. 42-45. Эхокардиографические предикторы неблагоприятного исхода операции тромбэндартерэктомии из ветвей лёгочной артерии.

В разделе «Материал и методы» авторы ничего не сообщают об использованных ими методах статистического анализа. Тогда как в разделе «Результаты» приводятся различные выражения, полученные с помощью статистических методов. Например, в табл. 2 авторы приводят результаты сравнения двух групп пациентов. Однако при этом не сообщают, с помощью каких статистических методов и критериев они проводили сравнение двух групп пациентов. Не сообщают авторы и о том, какие выборочные параметры представлены в статье в выражениях типа «88 ± 29,5» (см. табл. 2 слева). Можно предположить, что первая величина – среднее, а относительно второй величины стоящей после знака « ± » можно предположить, что это либо стандартное отклонение, либо ошибка среднего.

Таблица 2

Данные трансторакальной эхокардиографии у пациентов обеих групп

Эхокардиографический критерий	Группа 1	Группа 2	p
Конечный диастолический объем, мл			
правого желудочка	88±29,5	138,5±22,3	0,006
левого желудочка	72±20,3	44,6±10,3	0,017
Соотношение объемов правого и левого желудочков	1,16±0,56	3,27±1,07	0,0003
Площадь правого предсердия, см^2	25,5±7,34	42,8±0,17	0,006
Фракция выброса правого желудочка, %	42,5±7,9	33,6±5,6	0,05
Давление в легочной артерии, мм рт. ст.	74,7±25,4	96,5±9	0,089

В качестве положительного момента данной публикации можно отметить тот факт, что авторы приводят в статье достигнутые уровни значимости для использованных (но не упомянутых в разделе «Материал и методы») статистических критериев. Учитывая важность и актуальность данного исследования, целесообразно было при проведении статистического анализа полученных данных использовать такие методы многомерной статистики, как логистическая регрессия [44] и многомерная нелинейная регрессия.

Таблица 1

Показатели невротических изменений у пациентов с эрозивно-язвенными изменениями гастродуоденальной зоны перед операцией на сердце в условиях ИК (группа 1) и контрольной группы (группа 2)

Наименования шкал	Группа 1	Группа 2	Критерий t	p
Тревога	1,24±3,68^*	0,99±3,69^*	0,52	0,61
Депрессия	0,62±4,06^*	1,44±3,44^	-0,95	0,34
Истерическое реагирование	-0,09±3,32^*	0,51±3,11^*	-0,83	0,41
Астения	1,46±3,93^*	0,91±4,18^*	0,61	0,54
Обсессивно-фобические нарушения	1,20±3,01^*	0,86±2,79^*	0,52	0,61
Вегетативные нарушения	-0,38±7,49^*	-0,02±5,85^*	-0,23	0,82

средние значения находятся в зоне патологических изменений; ^ различия статистически недостоверны

Таблица 2

Показатели невротических изменений у геликобактер-позитивных (HP±) и геликобактер-негативных (HP-) пациентов с эрозивно-язвенными изменениями гастродуоденальной группы 1

Наименования шкал	HP±	HP-	Критерий t	p
Тревога	1,74±3,60^	-1,25±3,18^*	2,44	0,16
Депрессия	0,90±3,98^*	-0,82±4,36^*	1,23	0,22
Истерическое реагирование	0,61±2,88*	-3,63±3,27*	4,16	<0,001
Астения	1,46±3,93	-0,84±2,54*	2,06	0,042
Обсессивно-фобические нарушения	1,58±3,00	-0,69±2,37*	2,25	0,036
Вегетативные нарушения	0,62±7,37*	-5,39±6,19*	2,41	0,019

средние значения находятся в зоне патологических изменений; ^ различия статистически недостоверны

22) ПКК, № 2, 2010, с. 46-51. Этиологическая роль стресса, Helicobacter pylori и применения аспирина в генезе образования эрозий и язв гастродуоденальной зоны у пациентов на этапе подготовки к операции на сердце в условиях искусственного кровообращения.

В разделе «Материал и методы» авторы пишут: «Группы достоверно не различались между собой по половому составу (критерий $\chi 2$=0,064; p=0,80), возрасту (критерий Стьюдента t=0,258;

p=0,80), структуре сердечно-сосудистой патологии (критерий χ^2=0,056; p=0,81)» … Испытуемым предлагалось ответить на 68 вопросов, оценивая своё состояние по пятибалльной системе: 5 баллов – никогда не было, 4 балла – редко, 3 балла – иногда, 2 балла – часто, 1 балл – постоянно или всегда. Однако при этом авторы не проверили ограничения на использование критерия Стьюдента, что уже делает результаты, приведённые в табл. 1-2 сомнительными. В статье используются некорректные выражения, включающие в себя слово «достоверно» («Группы достоверно не различались», «различия не являлись статистически достоверными», «отсутствие достоверных различий» и т.д.). В своей статье «ЛОШАДЕНДУС СВАЛЕНДУС С МОСТЕНДУС» [70] мы рассматривали детально пример, в котором диссертант Омской государственной академии Леонтьев В.В. выполнил факторный анализ используя дискретные признаки. Тогда как данный метод применим лишь к непрерывным, количественных признакам. Подобно этому диссертанту авторы статьи использовали t-критерий Стьюдента применительно к 68 балльным признакам, которые имеют всего лишь 5 градаций. Т.е. такие дискретные признаки уже в силу своей природы не могут иметь нормального распределения. И по этой причине к ним нельзя применять t-критерий Стьюдента. Всё это говорит о том, что уровень знаний в области биостатистики среди исследователей Омской государственной медицинской академии весьма низок. В то же время, тот факт, что столь грубое нарушение корректности использования t-критерия Стьюдента осталось незамеченным редакционной коллегией журнала ПКК в составе 21 человека, а также рецензентами (если рукопись данной статьи рецензировалась), говорит также и об их уровне знаний по биостатистике. Сомнительные результаты использования t-критерия Стьюдента к балльным признакам авторы приводят в табл. 1-2.

Учитывая дискретную природу анализируемых признаков, авторам статьи вместо t-критерия Стьюдента целесообразно было применять непараметрические методы сравнения групп, анализ многомерных таблиц сопряжённости, анализ соответствий, лог-линейный анализ, и логистическую регрессию [44]. Очевидно, что сомнительные выводы, полученные в результате некорректного использования статистических методов, столь же сомнительно способны при их использовании в лечебной практике вызвать желаемую реакцию. Если, конечно, основной целью подобных исследований является получение полезной информации для её использования во врачебной практике, а не просто для использования её в публикации.

23) ПКК, № 1, 2011, с. 35-40. Влияние умеренной гипотермии на сывороточный уровень нейронспецифических белков, кислородное обеспечение и нейрокогнитивный статус пациентов при операциях реваскуляризации миокарда.

«При использовании тестов Шапиро – Вилка и Колмогорова – Смирнова выявлено, что распределение данных в совокупности носит **параметрический характер**. Сравнения групп с нормальным распределением проводились с помощью t – теста для двух независимых выборок. Анализ выборок при повторных измерениях осуществлялся с помощью критерия Фишера. … Статистически значимыми считались различия данных при p<0,05».

Положительным моментом описания является указание критического значения достигнутого уровня значимости p=0,05. Однако в тексте статьи встречаем выражения вида «достоверное отличие», что является некорректным оборотом [24]. Далее, авторы не поясняют, что означает оборот «параметрический характер». Если сделать в Яндексе поиск на словосочетание «параметрический характер», то обнаружим всего лишь несколько ссылок, в которых встречается данное словосочетание. Но ни в одной из них не анализируются распределения вероятностей. Вот некоторые предложения из этих ссылок. «При этом было установлено, что из-за нелинейности дифференциального уравнения (2) возбуждение носит не

только кинематический, но и **параметрический характер**». (И. И. Вульфсон. Колебания механизмов с гибкими связями при параметрических импульсах. Теория Механизмов и Машин. 2010. №2. Том 8. http://tmm.spbstu.ru/16/vulfson_16.pdf). «В случае параметрического синтеза САУ, алгоритм решения задачи оптимизации имеет **параметрический характер**». (Принципы автоматизации Систем Автоматического Управления. Методы и Алгоритмы. Студент гр. САПР-1В-96 Царев Александр В. http://radiolubite.narod.ru/referats/47/1061-013/index.html). «Проблема носит **параметрический характер**, если она может быть устранена изменением только параметров организационно-производственной системы» (Разработка и реализация управленческих решений. http://elearn.oknemuan.ru/?p=1&id=8).

Если говорить об использовании тестов Шапиро-Вилка (Шапиро-Уилка) и Колмогорова-Смирнова, то данные критерии могут быть использованы для проверки гипотез о разных законах распределения вероятностей. Судя по контексту, авторы проверяли гипотезу о нормальном законе распределения. Но и в этом случае выражение «**параметрический характер**» не может заменить собой конкретизацию закона распределения, поскольку ВСЕ законы распределения имеют разные наборы ПАРАМЕТРОВ. В том числе и те законы, которые не относятся к семейству нормальных. Для нормального закона это 2 параметра: генеральное среднее и генеральная дисперсия. Для всех других законов эти наборы будут включать разные параметры. Так, распределение Стьюдента имеет один параметр: f – число степеней свободы. Распределение Фишера имеет 2 параметра: f_1 и f_2 – степени свободы. Распределение Рэлея имеет один параметр – a, и т.д. Поэтому упоминая тесты (критерии) проверки гипотезы о законе распределения, необходимо конкретизировать название этого закона. Однако для корректного применения критерия Стьюдента при проверке гипотез о равенстве 2-х групповых средних, недостаточно только проверки нормальности. Необходимо выполнение и второго условия: равенства двух генеральных дисперсий (см. статью «Когда нельзя, но очень хочется, или Ещё раз о критерии Стьюдента» [32]. О проверке этого второго условия авторы не сообщают, из чего можно сделать вывод о том, что эта проверка не проводилась. Учитывая, что оба условия одновременно выполняются крайне редко, примерно в 2-5% случаев, можно утверждать, что результаты, полученные авторами с помощью критерия Стьюдента, не могут быть признаны надёжными.

24) ПКК, №1, 2011, с. 49-52. Нужна ли кардиоплегия при операции прямой реваскуляризации миокарда?

«Полученные данные обработаны с использованием программы статистического анализа StatPlus 2007 Professional. Результаты представлены в виде M ± m, где M – среднее значение, а m – стандартное отклонение». Положительным моментом описания является конкретизация выражения M ± m. Однако при этом авторы не сообщают, какие конкретно статистические методы и критерии были ими использованы. Между тем в программе StatPlus 2007 Professional имеется несколько десятков методов и критериев. Причём достаточно современных и глубоких методов анализа. Тогда как авторы статьи ограничились лишь самыми простейшими – сравнением средних в группах.

25) ПКК, № 3, 2010, с. 40-44. Оптимизация защиты головного мозга путем применения методики гипоксического прекондиционирования и анестезии ксеноном у кардиохирургических пациентов.

В таблицах 2-4 авторы приводят около 100 выражений вида 12,60 ± 2,78, однако при этом не указывают, какие конкретно величины связаны знаком « ± » . К недостаткам данной статьи

также следует отнести использование авторами выражений с использованием слова «достоверно» [24].

26) ПКК, № 3, 2010, с. 50-55. Проблема церебральных микроэмболических осложнений у кардиохирургических больных и гемореологические методы их профилактики.

В таблицах 2-4 авторы приводят около 100 выражений вида 9,8 ± 2,7, однако при этом не указывают, какие конкретно величины связаны знаком « ± ». К недостаткам данной статьи следует отнести использование авторами выражений с использованием слова «достоверно» [24]. В статье приводятся многочисленные результаты сравнений более 2 групп, с примечаниями типа «$p < 0,05$», однако авторы не сообщают, какие гипотезы и с помощью каких критериев они проверяли.

27) ПКК, № 4, 2010, с. 12-16. Сравнительная оценка вариантов хирургической коррекции аномалии Эбштейна.

В таблицах 2-4 авторы приводят более 180 выражений вида 10,6 ± 7,9, однако при этом не указывают, какие конкретно величины связаны знаком « ± » . К недостаткам данной статьи следует отнести использование авторами выражений с использованием слова «достоверно». [24] Проводя сравнение групп, авторы умалчивают, какими конкретно статистическими методами и критериями они производили данные операции. Учитывая, что в работе анализируются 7 групп сравнения, необходимо было учитывать проблему множественных сравнений, для чего нужно было значительно снизить критический уровень значимости. Однако, авторы видимо не подозревают о существовании такой проблемы, и потому не уменьшили принятый ими критический уровень значимости в 5%. Учитывая тот факт, что в работе использовались 9 количественных признаков, а также имелось немало и качественных показателей, описывающих анамнез и особенности проводимых операций, целесообразно было использовать в исследовании многомерные методы статистики, в частности, непараметрический дискриминантный анализ и логистическую регрессию [44].

28) ПКК, № 4, 2010, с. 17-22. Непосредственные результаты протезирования аортального клапана у пациентов с аортальным стенозом и систолической дисфункцией левого желудочка.

В разделе «Материал и методы авторы пишут: «Данные представлены в виде среднее ± стандартное отклонение (M ± m). Сравнение внутри и межгрупповых различий количественных признаков оценивали с помощью парного и непарного коэффициента Стьюдента, качественных – χ^2 Пирсона. Факторы риска госпитальной летальности определяли методом многофакторной логистической регрессии. Уровень значимости p принимали менее 0,05». К положительным моментам данной публикации следует отнести конкретизацию авторами выражений вида M ± m, а также указание критического уровня статистической значимости. К недостаткам же данной публикации следует отнести отсутствие проверки двух условий применимости критерия Стьюдента. С учётом того, что одновременное выполнение этих условий реализуется крайне редко, полученные с помощью критерия Стьюдента выводы авторов нельзя считать

доказательными. К недостаткам следует также отнести использование авторами выражений с использованием слова «достоверно» [24]. Авторы сообщают: «В нашем исследовании единственным независимым фактором риска возникновения госпитальной летальности, определённым методом многофакторной логистической регрессии, было время окклюзии аорты больше 160 мин». Между тем, в таблице 2 «Структура хирургических вмешательств, тип и размер протеза, время искусственного кровообращения и окклюзии аорты (p = NS)» приведённый признак «Время окклюзии аорты, мин» не имеет дискретных значений, содержащих градацию «Более 160 мин». Это непрерывный, количественный признак, имеющий для первой группы (n=22) среднее значение 120,8 ± 41,9, и для второй группы (n=19) среднее значение 126,6 ± 35,9. Поэтому непонятно, каким образом непрерывный количественный признак «Время окклюзии аорты, мин» был превращён с помощью метода логистической регрессии в дискретный признак. Возможно, что данное граничное значение представляет собой точку деления, оценённую с помощью иных методов анализа. Например, с помощью ROC-анализа. Но в этом случае необходимо объяснить процедуру оценки порогового значения. Поскольку существуют разные методы таких оценок. Напомним, что метод логистической регрессии [44] тем и привлекателен, что он позволяет использовать в клинических исследованиях для оценки взаимосвязи между дискретным, качественным признаком (в данном случае это признак с двумя градациями «Выжил – Умер») как непрерывные, количественные признаки, например, «Возраст, годы», «Время окклюзии аорты, мин» и т.д., так и дискретные, качественные признаки, например, «Пол», «ХОБЛ», «ХСН», и т.д. Поэтому в том случае, когда алгоритм оценки уравнения логистической регрессии отбирает признак «Время окклюзии аорты, мин», то это не означает, что данный признак становится дискретным. В своей статье авторы не сообщают также никаких параметров полученного ими уравнения, содержащего упоминаемый ими единственный предиктор «время окклюзии аорты больше 160 мин». В частности, отсутствуют такие параметры этого уравнения, как коэффициенты уравнения, достигнутые уровни значимости этих коэффициентов, отношение шансов для установленного предиктора. Нет и такого важного показателя для оценки адекватности полученного уравнения, как результаты проверки уравнения с помощью критерия Хосмера-Лемешова. Нет и не менее важного показателя – результата проверки однородности распределения остатков с помощь критерия Пирсона Хи-квадрат.

И, наконец, нет самого важного показателя качества такого уравнения, показателя конкордации – показателя согласия фактических значений зависимого признака «Летальность» (Выжил – Умер) и расчётных. Когда показатель конкордации близок к 1, то это означает высокую степень связи зависимого признака «Летальность» и набора предикторов. И наоборот, если этот показатель близок к 0,5, или менее 0,5, то данное уравнение не имеет никакой ценности [44]. Кроме того, учитывая что в данном исследовании имеется 15 потенциальных предикторов, приведённых в таблице 2 «Структура хирургических вмешательств, тип и размер протеза, время искусственного кровообращения и окклюзии аорты (p = NS)», вызывает сомнение, что из этих 15 признаков в уравнение логистической регрессии алгоритмом был отобран лишь 1 предиктор. Весьма вероятно, что такой результат объясняется отсутствием у исполнителя статистического анализа этих данных необходимых знаний и опыта оценки логистических уравнений. Такая версия вполне реальна, поскольку данный исполнитель даже не указал в разделе «Материал и методы» ни

название пакета, в котором проводился данный анализ, не привёл ни одной ссылки на литературные источники с описанием этого метода, и не указал, какой конкретно алгоритм оценки уравнения (а их очень много) был им использован. Отметим также, что в тех случаях, когда подобный анализ выполняется силами профессиональных биостатистиков, используются различные алгоритмы оценок уравнений логистической регрессии. В результате чего получается не единственное, уникальное уравнение, а набор различных уравнений. Из которых далее отбирается одно, или несколько, наиболее ценных и полезных уравнений. Так автор этих строк за последние 5 лет получил более 5 тысяч оценок уравнений для исследователей в разных областях медицины, однако из них для реального использования было отобрано лишь порядка 200-300.

В статье авторы приводят ссылку под номером 18:

18. Morris J.J., Schaff H.V., Mullany C.J. et al. // Ann. Thorac. Surg. 1993. V. 56. P. 22–29.

Ниже приводится полное название этой статьи:

Determinants of Survival and Recovery of Left Ventricular Function After Aortic Valve Replacement

James J. Morris, MD, Hartzell V. Schaff, MD, Charles J. Mullany, MD, Amita Rastogi, MD, Christopher G. A. McGregor, MD, Richard C. Daly, MD, Robert L. Frye, MD, and Thomas A. Orszulak, MD

Division of Thoracic and Cardiovascular Surgery and Division of Cardiovascular Diseases and Internal Medicine, Mayo Clinic and Mayo Foundation, Rochester, Minnesota

В этой статье (*Determinants of Survival ...)* авторы весьма подробно описывают использованные ими статистические методы анализа своих наблюдений. И такое описание могло быть стать примером для подражания цитирующим эту статью. Ниже приводим копию раздела «Анализ данных» из этой статьи.

«Data Analysis
Baseline patient variables used in this analysis and tested for association with risk were as follows:
Preoperative variables
Age
Sex
Body surface area
New York Heart Association class
Status of operation
Elective
Urgent
Emergency
None
Right bundle-branch block
Left bundle-branch block
Second- or third-degree heart block
Type of valvular heart disease
Stenosis
Insufficiency
Combined stenosis and insufficiency
Conduction defect
Left ventricular ejection fraction

Left ventricular end-diastolic pressure
Aortic valve gradient
Extent of coronary artery disease (≥50% stenosis)
None
1-vessel disease
2-vessel disease
3-vessel disease
Left main disease (≥50% stenosis)
Operative variables
Year of operation
Prosthetic valve type
Mechanical
Bioprosthetic
Homograft
Prosthetic valve size
Completeness of revascularization (calculated as
number of major coronary vessels bypassed divided
by number of major vessels diseased [≥50% stenosis]; = 0 if no vessels diseased)

These variables are similar to risk factors previously reported [4, 7, 9-11]. Differences between patient groups in baseline clinical, angiographic, echocardiographic, operative, and follow-up variables were tested for statistical significance by *t* or χ^2 tests or by analysis of variance as appropriate. **Multivariate analysis of all risk factors relating to outcome was performed for the overall population by using a *stepwise Cox proportional hazards model* [12] for overall survival and a regression analysis for postoperative change in ejection fraction.** Variables were included in multivariate analyses if univariate significance was less than or equal to 0.10. Unadjusted overall patient group survival probabilities relative to baseline characteristics were compared by a log-rank test. All values are mean ± standard deviation except as otherwise noted. A *p* value less than 0.05 was considered significant ».

Table 4. Multivariate Analysis of Risk Factors for Mortality in All Patients

Variable	χ^2	*p* Value
Age	18	0.0001
Ejection fraction	16	0.0001
Extent of coronary artery disease	8	0.006
Prosthetic valve size	5	0.03
NYHA class	4	0.04

NYHA = New York Heart Association.

Table 6. Multivariate Analysis of Risk Factors for Mortality Including Postoperative Change in Ejection Fraction

Variable	χ^2	*p* Value
Preoperative ejection fraction	31	0.0001
Postoperative change in ejection fraction	21	0.0001
Age	13	0.0003
Extent of coronary artery disease	8	0.009

Данное описание состоит из 262 слов, и содержит не только перечисление анализируемых признаков, но и указание на то, что при анализе был ***использован пошаговый алгоритм для регрессии Кокса***. При этом авторы дают ссылку на источник [12], в котором как раз и описывается модель Кокса:

12. Cox DR. Regression models and life tables. J Res Stat SOC [B] 1972;34:187-201.

Слева приводим табл. 4 и 6 из этой статьи (*Determinants of Survival ...)*, в которых видно, что в качестве предикторов пошаговый алгоритм отобрал из всего перечня потенциальных предикторов более 1 признака. При этом для каждого предиктора приводится как значение

критерия Хи-квадрат, так и значение достигнутого уровня статистической значимости. Разумеется, одной из причин отбора нескольких предикторов может быть тот факт, что авторы анализировали результаты наблюдений за 1012 пациентами. Тогда как в статье «Непосредственные результаты протезирования аортального клапана у пациентов с аортальным стенозом и систолической дисфункцией левого желудочка» из ПКК анализируются наблюдения лишь за 41 пациентом. Однако и в таких случаях использование различных алгоритмов оценки уравнения логистической регрессии даёт возможность получить более 1 уравнения. И это тем более возможно, поскольку авторы статьи оперировали несколькими десятками показателей. Детальный же анализ и сравнение таких уравнений позволяет, в конечном счёте, получить наиболее оптимальные и ценные уравнения. Разумеется, для реализации этого варианта требуется участие опытного биостатистика.

29) ПКК, № 4, 2010, с. 23-28. Протезирование митрального клапана биологическими протезами «КемКор» и «ПериКор»: отдалённые результаты.

В разделе «Материал и методы авторы пишут: «Статистический анализ проводили с помощью программы Statistica 7.0. Анализ взаимосвязей признаков проводили **методами** нелинейного многомерного моделирования. При анализе выживаемости использовали метод Каплана-Мейера, а также таблицы выживаемости. Уровень значимости p, при котором отвергалась нулевая гипотеза, принимали равным 0,05». В статье авторы используют порядка сотни выражений вида «50 ± 12», однако при этом не сообщают в разделе «Материал и методы» какие именно величины связаны символом ± . Если первое число видимо является средней величиной, то второе число может быть как стандартным отклонением, так и ошибкой среднего. Ещё одна неясность касается авторского утверждения «Анализ взаимосвязей признаков проводили методами нелинейного многомерного моделирования». Из этой фразы совершенно непонятно, о каких конкретно методах моделирования говорят авторы. Если они подразумевают под нелинейными методами метод Каплана-Мейера, то он априори является нелинейным. Кроме того, речь идёт не о единственном нелинейном методе, а о многих **методах,** т.е. о неких иных методах моделирования. Однако в тексте статьи данные модели не приводятся. Положительным моментом статьи является указание авторами критического уровня статистической значимости.

30) ПКК, № 4, 2010, с. 29-35. Клинико-ангиографическая эффективность использования покрытых баллонных катетеров у больных с рестенозом ранее имплантированных коронарных стентов.

В разделе «Материал и методы авторы вообще никак не освещают статистические аспекты своего исследования. В частности, не объясняются, какие именно величины связаны символом « ± » в выражениях вида 58,6 ± 8,02. В пояснениях к табл. 2 «Данные ЭхоКГ больных двух групп, авторы приводят выражение «p > 0,05». Однако умалчивают, какие конкретно статистические гипотезы проверялись при этом, и какими конкретно статистическими критериями.

31) ПКК, № 4, 2010, с. 36-41. Клинико-лабораторная оценка липидтранспортной функции крови у пациентов с острым инфарктом миокарда с подъёмом сегмента ST.

В разделе «Материал и методы» авторы пишут: «Статистическую обработку проводили с помощью **непараметрических данных**». Уважаемые читатели! Вдумайтесь в смысл этой фразы. Получается, что вначале 8 авторов статьи собрали исходные данные о состоянии пациентов. Затем

к ним добавили (или применили ?) некие «непараметрические данные», и с помощью этих «непараметрических данных» стали анализировать данные о состоянии пациентов? Именно одним предложением из 6 слов 8 авторов и ограничивают своё описание статистических аспектов своего исследования. Уже один этот факт говорит об отношении авторов к использованному ими статистическому инструментарию и к выводам, полученным с его помощью. Как видим, авторы скрыли от читателей и редколлегии журнала многие аспекты этого исследования. Например, они не объяснили читателям и редколлегии журнала, что за величины используются ими в выражениях вида «0,947±0,06». В тексте статьи приводятся 3 рисунка и 2 таблицы, в которых использованы выражения «$p < 0,05$». Это свидетельствует о том, что авторы проверяли некие статистические гипотезы. Однако при этом авторы скрыли от читателей и редколлегии журнала, какими конкретно статистическими методами и критериями проверялись эти гипотезы. Также не объяснено, что они подразумевают под выражением «**непараметрические данные**». А такое объяснение необходимо, поскольку не существует «непараметрических данных». **Есть непараметрические методы статистики, а непараметрических данных нет, и не может быть в принципе**. Чтобы пояснить эту мысль, обратимся к толкованию слова «параметр». Вот как объясняется этот термин в «Словаре иностранных слов» (М.: Изд-во «Русский язык», 1989. – 624 с .) на стр. 370: «[<гр. parametrön отмеривающий] – 1) *мат.* величина, входящая в формулы и выражения, значения которой является постоянным в пределах рассматриваемой задачи». Аналогичные толкования этого термина можно найти во многих других словарях. Если же обратиться к Энциклопедии «Викисловарь» (URL: http://ru.wiktionary.org/wiki/параметр), то вот как там предельно сжато описано значение этого слова применительно к математике: «количественный показатель, характеризующий важное свойство чего-либо». Если, исходя из этого объяснения, интерпретировать выражение «непараметрические данные», то получается что данные, которые анализировались авторами, не имеют вообще никаких параметров, т.е. нет ни средних величин, ни минимальных, ни максимальных, ни стандартных отклонений, ни объёма выборки и т.д. А ведь это и есть те самые «параметры» анализируемых данных. Авторы, использующие выражение «непараметрические данные», не являются начинающими исследователями, вообще не имеющими никакого опыта использования статистики в медицине. Это 2 доктора медицинских наук, профессора, один из них даже академик РАМН, директор НИИ, другой – заведующий отделом этого НИИ и зав. кафедрой КГМА; 3 кандидата медицинских наук работающих в этом самом НИИ, один из них заведующий лабораторией, и 3 младших научных сотрудника упомянутого НИИ. Использование некорректного выражения «непараметрические данные» означает, что авторы даже не знакомы с такими понятиями, как «параметрические методы статистики» и «непараметрические методы статистики». Напомним, что подобные статистические нелепости были нами детально проанализированы в статье «ЛОШАДЕНДУС СВАЛЕНДУС С МОСТЕНДУС» [70].

Далее авторы пишут: «У пациентов с рестенозами выявлен повышенный уровень ЛП(а), при этом **достоверных различий** по другим измерявшимся показателям, а именно ЛПВП, ЛПНП, ЛПОНП, АпоА-1 и Апо В, не обнаружено. … Вместе с тем результаты недавних исследований Г.Г. Арабидзе и др. [2009] показали отсутствие **достоверной взаимосвязи** показателя ЛП(а) с заболеваемостью ишемической болезнью сердца, а имеющаяся некоторая взаимосвязь ЛП(а) с известными факторами риска ИБС – артериальной гипертонией, диабетом, а также возрастом и полом, по мнению авторов, обладает небольшим уровнем значимости [2]. … Вместе с тем на фоне увеличения содержания атерогенных липопротеинов и апопротеина В выявлены достоверное снижение антиатерогенного ХС ЛПВП и тенденция к уменьшению входящего в их состав апопротеина А (рис. 1) ... При этом у мужчин эти изменения выражены в большей степени, а в группе женщин к 12-м суткам наблюдения отмечается **достоверное повышение** апо-В и

тенденция к увеличению остальных изучаемых параметров (рис. 3). … у пациентов с острым коронарным синдромом в первые сутки содержание ЛП(а) **достоверно повышается**… ». Использование подобных выражений со словом «**достоверно**», с точки зрения статистики, является некорректным,[24] и свидетельствует о незнании авторами статьи базовых понятий и терминологии статистики [65]. Детальный анализ этой некорректности дан в статье известного специалиста по доказательной медицине Н.А. Зорина «О неправильном употреблении термина "достоверность" в российских научных психиатрических и общемедицинских статьях»[24]. Обратимся к энциклопедии «Вероятность и математическая статистика» (Научное издательство «Большая российская энциклопедия», М.: 1999. – 910 с.) На странице 178 этой энциклопедии прочитаем: «… понятие «Достоверность» связано с ***достоверными событиями***, происходящими с вероятностью 1». Применительно же к результатам статистического анализа выборочных наблюдений, вероятности этих событий, формулируемых в виде статистических гипотез, не равны 1. И в силу этого подобные обороты в данном контексте являются неверными. Использование авторами таких оборотов тем более нелогично, что в этой же самой статье авторы используют достаточно часто и правильные выражения вида «статистически значимые различия». Наиболее вероятная причина такого противоречия может заключаться в том, что при наличии 8 соавторов отдельные фрагменты текста могли готовиться разными соавторами. И в силу этого использовались различные формулировки. При этом, как известно, часть соавторов часто добавляется с целью скорейшей публикации статьи. Например, аспиранту необходимо в определённые сроки до защиты диссертации опубликовать определённое количество статей. И чтобы быстрее их опубликовать, в соавторы включают как своё непосредственное начальство, так и директоров НИИ или ректоров вузов. Полагаю, что данная технология знакома многим… Как известно, количество авторов публикаций и цитируемость этих публикаций имеет отрицательную корреляцию. Т.е. чем больше количество соавторов, тем реже цитируются эти публикации. Возможной причиной этого как раз и являются подобные конфликты содержания отдельных фрагментов текста, написанные разными соавторами. К недостаткам данной публикации следует отнести также и то, что анализируя более 2 групп сравнения, авторы не учитывают проблему множественных сравнений.

32) ПКК, № 4, 2010, с. 42-45. Гемодинамика и гидродинамический статус при системном воспалительном ответе инфекционной и неинфекционной этиологии на фоне продолжительной заместительной почечной терапии.

По сравнению с предыдущей статьёй (опубликованной в этом же номере ПКК на стр. 36-41) в данной публикации авторы дали несколько более подробное описание статистических аспектов своего исследования. Напомню, что в предыдущей статье всё описание заняло 1 предложение из 6 слов. В анализируемой же статье это описание включало уже 3 (NB!) предложения, и состояло из 33(NB!) слов: «Все данные в исследовании представлены как среднее арифметическое ± стандартное отклонение (M ± σ). Статистически значимыми различиями при сравнении качественных данных считали при уровне $p<0,05$. Оценка характера распределения проводилась по критерию Колмогорова – Смирнова». Итак, положительный момент данного описания – конкретизация выражения «арифметическое ± стандартное отклонение (M ± σ)». Однако выражение M ± σ нуждается в корректировке. В статистике принято обозначать греческими буквами генеральные параметры, т.е. популяционные параметры. Тогда как выборочные параметры, вычисляемые по имеющимся выборкам, принято обозначать их латинскими аналогами. Например, генеральное среднее обозначают буквой «μ», а выборочное среднее буквой «M». Аналогично генеральное стандартное отклонение обозначают буквой «σ», а выборочное – «s». Таким образом, более правильно было бы вместо «M ± σ» написать «M ± s».

Прочитай и перешли своим коллегам…

Тем не менее, сам факт использования выражения «арифметическое ± стандартное отклонение», которое отсутствовало в предыдущей статье, уже говорит о многом. В частности, если сравнить порядок расположения авторов данной и предыдущей статей, то можно предположить, чем вызвано появление этого выражения. Как видим, в описании статистики в данной статье отсутствует и выражение «**непараметрические данные**». Напротив, появилось упоминание об оценке вида распределения с помощью критерия Колмогорова-Смирнова. К сожалению, одного лишь упоминания этого критерия недостаточно. Ведь авторы не сообщают читателям и редакционной коллегии, на соответствие какому конкретно закону распределения проводилась эта оценка. Можно предположить, что оценивалось согласие с нормальным распределением. Но и в этом случае также есть тонкости использования данного критерия, которые не упоминаются авторами.

В исследовании анализировались изменения 12 количественных признаков в 4 периода наблюдения: «Исходно», «1-е сутки», «2-е сутки» и «3-е сутки». Таким образом, имеется один фактор с 4 градациями. В свою очередь каждая из этих 4-х групп разбита на 2 подгруппы: «Позднее начало ЗПТ» и «Раннее начало ЗПТ» (ЗПТ – заместительная почечная терапия). Помимо этого имеется и ещё один дополнительный фактор с 2 градациями: «СВО+инфекция» и «СВО+И». Таким образом, налицо 3 фактора влияющих на значения 12 анализируемых 12 количественных признаков. В таблицах 1 и 2 авторы приводят результаты сравнения около 200 средних. При этом в примечаниях к таблицам используются выражения вида «$p < 0,05$». Однако авторы скрыли от читателей и редакционной коллегии, какими конкретно статистическими методами и критериями они пользовались для сравнения групп. Между тем, это очень важный аспект подобного исследования. Обозначим имеющиеся 3 фактора буквами A, B и C. Анализируя характер изменения каждого из 12 количественных признаков возможно попытаться установить степень влияния каждого из этих трёх факторов. Обозначим степень такого влияния как η^2_A, η^2_B, η^2_C. Помимо индивидуальных влияний этих факторов в отдельности, возможны также проявления и так называемых эффектов взаимодействия. Это 3 двойных эффекта взаимодействия, и один тройной эффект: η^2_{AB}, η^2_{AC}, η^2_{BC}, η^2_{ABC}. Суть этих эффектов заключается в том, что для некоторых комбинаций двух или трёх факторов средние групповые значения не будут различаться, тогда как для оставшихся комбинаций могут наблюдаться существенные, значимые изменения. Фактически мы имеем классический пример данных для анализа с помощью многофакторного дисперсионного анализа.

Учитывая тот факт, что в исследовании анализировалось 12 количественных признаков, целесообразно было использовать и несколько многомерных методов статистического анализа. Например, дискриминантный и факторный анализ, а также анализ канонических корреляций. Кроме того, учитывая, что помимо этих 12 количественных переменных в анамнезе пациентов фиксировалось немало и других признаков, в т. ч. качественных, целесообразно было использовать и такой мощный метод, как логистическая регрессия [44].

33) ПКК, № 4, 2010, с. 51-55. Распределение химических элементов таблицы Менделеева в сердечно-сосудистой системе кардиохирургических больных.

В разделе «Материал и методы» авторы пишут: «Статистическая обработка результатов проводилась с помощью программы Microsoft Excel 2000. **Достоверность различий средних величин** и корреляционных взаимоотношений проводили с помощью t-критерия Стьюдента. Достоверными считали различия при $p<0,005$». К недостаткам данной публикации следует отнести сомнительность полученных выводов, поскольку авторы не проверяли выполнение ограничений на использование критерия Стьюдента. Отметим, что одно из этих ограничительных условий в

принципе невозможно проверить, используя программу Microsoft Excel 2000. Далее, авторы не сочли необходимым довести до сведения читателей и редакционной коллегии, какие конкретно коэффициенты корреляции были ими оценены. Как известно, имеется довольно много различных видов корреляционных коэффициентов. Неясно и то, различия каких корреляционных взаимоотношений проводились авторами. Вызывает также вопрос и об использовании столь низкой величины критического уровня статистической значимости, который авторы задали равным 0,005. Что это? Опечатка, допущенная работниками журнала, или же сознательный выбор самих авторов статьи? Ведь в предыдущей аналогичной статье «Распределение химических элементов в сосудистой системе больных ишемической болезнью сердца» (ПКК, № 2, 2010, с. 22-25. См. выше) авторы использовали критический уровень значимости равный 5%. Авторы используют некорректные выражения со словом «достоверность». Учитывая тот факт, что в исследовании анализировалось распределение 17 химических элементов, целесообразно было использовать в данном исследовании не «плоский» подход, заключающийся в сравнении групповых средних, а применить методы многомерной статистики, такие, как факторный анализ, анализ канонических корреляций, дискриминантный анализ и логистическая регрессия [44]. **Учитывая огромную важность цели данного исследования, предлагаем авторам данной статьи бесплатное проведение детального статистического анализа собранных ими данных, при условии публикации в ПКК новых результатов, и их сравнении с опубликованными в данной статье.**

34) ПКК, № 3, 2011, с. 17-21. Отдалённые результаты хирургической процедуры Maze у пациентов с клапанной патологией и фибрилляцией предсердий.

В разделе «Материал и методы» авторы пишут: «При статистической обработке данных были использованы методы описательной статистики и анализа выживаемости (актуриальный метод таблиц жизни). Уровень значимости принят равным 0,05. Обработка данных производилась с помощью программы Statistica 6.0». Хорошо, что авторы указывают критический уровень значимости равный 5%. Положительный момент статьи также и в том, что в ней авторы не используют архаичные выражения вида «$p<0,05$» или «$p>0,05$», а приводят фактические, достигнутые значения уровней значимости. Например, пишут «$p = 0,038$». Однако при этом почему-то скрывают от редакции и читателей, какие конкретно статистические критерии они применяли. Между тем в тексте статьи немало приводится результатов, указывающих на проверки статистических гипотез этими неизвестными для читателей критериями и методами. Не указывают авторы и на то, какие конкретно величины приводятся ими в выражениях вида 39,0±17,7. Можно предположить, что первое число – это среднее. Тогда как вторым числом могут быть самые разные параметры. В аннотации к статье авторы пишут: «**Методом многофакторного анализа** выявлено, что выраженная дилатация левого предсердия и длительность мерцательной аритмии более 5 лет выступают основными предикторами возврата ФП». Но при этом авторы не уточняют, какой конкретно метод многофакторного анализа был ими использован. Такое уточнение обнаруживается далее в тексте статьи: «Для выделения предикторов возврата ФП на этапе динамического наблюдения был проведён многофакторный регрессионный анализ. Выраженная (более 6,5 см) атриомегалия ($p = 0,038$) и длительность (более 5 лет) фибрилляции предсердий ($p<0,001$) оказывают значительное влияние на этот показатель».

Итак, авторы сообщают, что использовали в своём исследовании многофакторный регрессионный анализ. Однако такая формулировка метода достаточно расплывчата, и не конкретизирует ни структуру полученных уравнений, ни алгоритмы, по которым оценивались эти

уравнения. Как известно, существуют различные виды множественной регрессии. К примеру, хронологически первым видом множественной регрессии была множественная линейная регрессия. В этом случае зависимая переменная представляет собой непрерывную количественную переменную, обозначаемую обычно как Y. Объясняющие же переменные (предикторы) также являются количественными, и обозначаются обычно как X_1, X_2, …. X_i. В этом случае уравнение множественной линейной регрессии выглядит следующим образом: $Y = A_0 + A_1*X_1 + A_2*X_2 + \ldots + A_i*X_i$, где величины A_i – регрессионные коэффициенты. Однако из фразы «Для выделения предикторов возврата ФП на этапе динамического наблюдения был проведён многофакторный регрессионный анализ. Выраженная (более 6,5 см) атриомегалия (p = 0,038) и длительность (более 5 лет) фибрилляции предсердий (p<0,001) оказывают значительное влияние на этот показатель» следует, что зависимой переменной является переменная «Возврат фибрилляции предсердий». Т.е. это качественная переменная, имеющая, как минимум, 2 градации: «Есть возврат фибрилляции предсердий» и «Нет возврата фибрилляции предсердий». Из чего следует, что авторы использовали нелинейную регрессию, вероятнее всего логистическую множественную регрессию [44]. В пользу этой версии говорит и тот факт, что в качестве предиктора авторы упомянули качественный признак «Выраженная (более 6,5 см) атриомегалия», т.е. фактически использовали не количественный показатель «Средний размер левого предсердия, см», а дискретный показатель «Степень атриомегалии». В пользу версии об использовании авторами логистической регрессии говорит и такая фраза: «При проведении многофакторного анализа причин возникновения инсульта наиболее статистически значимым (p<0,001) была ФП и размер левого предсердия более 6,0 см в отдалённом периоде (p = 0,021)». К сожалению, авторы не только не конкретизируют метод «многофакторного анализа», но также и не приводят наиболее важные показатели этого анализа, характеризующие степень надёжности полученных уравнений, таких как коэффициент детерминации R^2, показатель конкордации и т.д.

Ещё одним недостатком данной публикации следует считать использование авторами противоречащих друг другу выражений вида «статистически достоверное» и «статистически значимое», [24] свидетельствующее о том, что авторы не понимают статистического смысла двух этих утверждений.

35) ПКК, № 3, 2011, с. 39-46. Превентивная внутриаортальная баллонная контрпульсация или левосимендан? Что лучше у кардиохирургических пациентов высокого риска?

В разделе «Материал и методы» авторы пишут: «Для оценки нормальности распределения количественных признаков применялась визуальная оценка частотного распределения с последующей оценкой нормальности с использованием критериев Шапиро – Уилка и Д'Агостино. **Параметрические признаки** описаны в виде среднего значения и стандартного отклонения (в скобках). **Непараметрические количественные признаки** приведены в виде медианы и границ межквартильного интервала (в скобках). Для бинарных признаков приведена доля и 95%-й доверительный интервал для доли (в скобках). Сравнительный анализ количественных признаков выполнен с помощью дисперсионного анализа (при нормальном распределении признака; после проверки однородности дисперсии критерием Бартлетта) либо критерием Крускала – Уолиса. Последующие (post hoc) межгрупповые сравнения проведены с помощью критерия Тьюки – Крамера (для параметрических признаков) либо критерием Коновера. Сравнения качественных признаков проводились точным критерием Фишера – Фримена – Холтера с последующим поиском межгрупповых различий критерием Тьюки (после арксинусной трансформации долей). Корреляционный анализ выполнен с помощью критерия Спирмена. Результаты корреляционного

анализа представлены в виде коэффициента корреляции и (в скобках) его 95%-го доверительного интервала. Для определения точки разделения параметров строилась характеристическая кривая (ROC). Площадь под ROC (AUC) представлена в виде её значения, границ ДИ и вероятности отличия AUC от площади под диагональной линией. Значение на ROC с наибольшей суммой чувствительности и специфичности определялось как точка разделения параметра. Для точки разделения приведены значения чувствительности, специфичности и отношения шансов (ОШ) с 95%-ми доверительными интервалами. Многофакторный анализ для дихотомической переменной проводился с помощью пошаговой бинарной логистической регрессией. Независимые переменные вводились и исключались из анализа на основе критерия отношения правдоподобия ($p<0,05$) для включения и $p>0,1$ для исключения переменной). Для всех статистических критериев ошибка первого рода устанавливалась равной 0,05. Нулевая гипотеза (отсутствие различий) отвергалась, если вероятность (p) не превышала ошибку первого рода. Статистический анализ данных проведён согласно общепринятым методам с использованием лицензионной программы Stata 11.2 (StatCorp, США)».

Приведённое выше описание, включающее в себя 270 слов, впечатляет своей основательностью и детализацией. Однако сравнение этого подробного описания и текста самой статьи обнаруживает явные противоречия. В частности, настораживает отсутствие в тексте результатов применения описанных выше статистических методов. В итоге складывается впечатление, что описание статистических аспектов является неким «инородным телом» в данной статье. Т.е. вполне возможно, что описание статистических методов в разделе «Материал и методы» было написано одним автором, а все остальные разделы – другими. И эти противоречия не были обнаружены ни авторами, ни редакцией журнала. Наиболее показательным противоречием является утверждение «Для всех статистических критериев ошибка первого рода устанавливалась равной 0,05» и выражения в тексте статьи с использованием слова «достоверно» [24]. Читаем дальше. В разделе «Материал и методы» написано: «Корреляционный анализ выполнен с помощью критерия Спирмена. **Результаты корреляционного анализа представлены в виде коэффициента корреляции и (в скобках) его 95%-го доверительного интервала**». Однако поиск в статье этих самых коэффициентов корреляции и их доверительных интервалов, не даёт никаких результатов. Нет в тексте статьи ни коэффициентов корреляции, ни их доверительных интервалов. Возникает резонный вопрос: где же были «представлены» эти упомянутые коэффициенты корреляции и их доверительные интервалы? Быть может, они были представлены в рукописи статьи, но затем удалены редакцией ПКК?

Ещё больше удручает следующий фрагмент описания статистических аспектов данного исследования. «**Параметрические признаки** описаны в виде среднего значения и стандартного отклонения (в скобках). **Непараметрические количественные признаки** приведены в виде медианы и границ межквартильного интервала (в скобках). … Последующие (post hoc) межгрупповые сравнения проведены с помощью критерия Тьюки – Крамера (для **параметрических признаков**) либо критерием Коновера». Это описание говорит о том, что авторы статьи не понимают смысл такого термина, как «параметр», применительно к контексту

статьи. Есть непараметрические методы статистики, а **непараметрических признаков нет, и не может быть в принципе**. Выше мы уже рассматривали подобные некорректности.

В разделе «Материал и методы» авторы сообщают об использовании ROC-анализа (ROC-кривых). В частности, сообщается что «Площадь под ROC (AUC) представлена в виде её значения, границ ДИ и вероятности отличия AUC от площади под диагональной линией». Увы, опять же поиски в тексте статьи подобных ROC-кривых и AUC, и «вероятности отличия AUC от площади под диагональной линией» также не приносят никаких результатов. В разделах «Результаты» и «Обсуждение» ROC-кривые упоминаются лишь дважды. Вот эти цитаты. 1) «В результате проведения ROC-анализа **мы получили данные, свидетельствующие о прогностической значимости мозгового натрийуретического пептида**». 2) «В нашем исследовании в результате проведения ROC-анализа **мы получили данные, свидетельствующие о прогностической значимости мозгового натрийуретического пептида**».

Напомним, что при построении ROC-кривых можно использовать один единственный количественный признак, например, ЧСС или САД, и, изменяя его, оценивать такие показатели, как чувствительность и специфичность, и далее строить ROC-кривые. Поскольку авторы сообщают об использовании бинарной логистической регрессии, то в этом случае в качестве такой непрерывной переменной следует использовать так называемый beta-параметр [44]. Именно этот beta-параметр и соединяет в себе комбинации значений различных признаков, вошедших в качестве предикторов в уравнение логистической регрессии. Далее он используется в виде показателя степени $\mathbf{beta_{ij}}$ в выражении для оценки вероятности отнесения конкретного i-того наблюдения к конкретной j-той группе сравнения:

$$P_{ij} = \frac{e^{beta_{ij}}}{1 + e^{beta_{ij}}}$$

Таким образом, после прочтения статьи возникает резонный вопрос: с какой целью авторы статьи приводят в разделе «Материал и методы» столь подробное описание различных статистических методов, тогда как результаты их использования отсутствуют в разделах «Результаты» и «Обсуждение»? Отметим, что данная проблема асимметрии описания использованных статистических методов и результатов их использования не уникальна. Она достаточно часто встречается в публикациях. Так в нашем докладе [46] на международной конференции по доказательной медицине в Ереване, мы приводили аналогичные примеры по диссертациям, защищённым в ЦНИИ организации и информатизации Минздрава.

36) ПКК, № 3, 2011, с. 47-50. Субпопуляционный состав лимфоцитов после кардиохирургических вмешательств в условиях искусственного кровообращения.

В разделе «Материал и методы» авторы пишут: «Значимость различий **непараметрически распределённых** зависимых величин оценивали критерием Фридмана …». Приведённое выше выражение является некорректным, поскольку не бывает «непараметрических распределений» (вероятности). Все распределения вероятности имеют параметры.

Прочитай и перешли своим коллегам…

37) ПКК, № 3, 2011, с. 57-62. Оптимизация противоишемической защиты головного мозга во время операций на экстракраниальных артериях.

В разделе «Материал и методы» автор пишет: «Для оценки и анализа полученных данных применялись стандартные методы описательной статистики, **в связи с тем, что все данные имели распределения, отличные от нормального**». Совершенно абсурдная аргументация использования стандартных методов описательной статистики [65]. Из приведённого выше утверждения следует, что для признаков имеющих нормальное распределение, стандартные методы описательной статистики не могут быть использованы. Читаем дальше: «Для анализа связи между двумя признаками, **выявления различий в частоте** неблагоприятных клинических признаков применялся метод ранговой корреляции Спирмена». Непонятно, каким образом автор, используя ранговую корреляцию Спирмена, выявлял «различия в частоте»? Читаем далее: «Для оценки сопряжённости процессов использовали корреляционный анализ с определением коэффициентов достоверности корреляции …». В этом утверждении автор не уточняет, какие конкретно коэффициенты корреляции оценивались им. Ведь уже в предыдущей фразе он сообщал об использовании ранговой корреляции Спирмена. Не сообщает автор и о том, что такое «**коэффициенты достоверности**»? Читаем далее: «Различия считали достоверными при $p<0,05$». Данное выражение является некорректным. Учитывая, что автор является заочным аспирантом кафедры анестезиологии и реаниматологии Кемеровской государственной медицинской академии Росздрава, можно предположить, что такой уровень знаний по статистике он получил именно на этой кафедре.

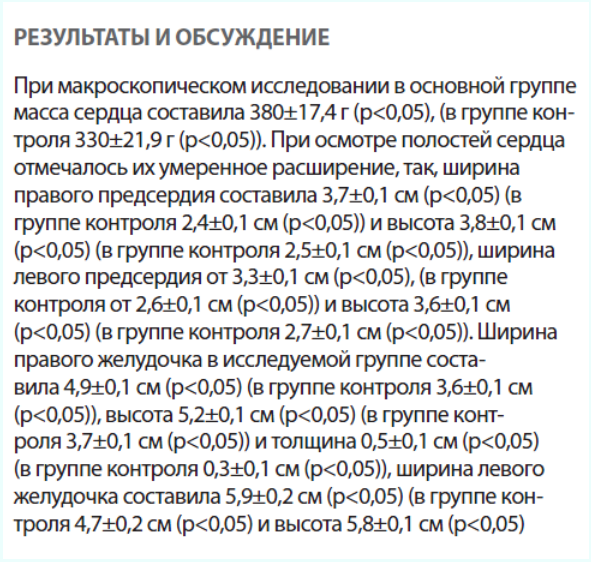
38) ПКК, № 3, 2011, с. 63-66. Полиморфизм генов ФНО-α, ИЛ-1β , iNOS и особенности системной воспалительной реакции у больных с хронической сердечной недостаточностью.

В разделе «Материал и методы» авторы пишут: «Сравнение средних значений анализируемых показателей проводили с помощью t-критерия Стьюдента или U-критерия Манна – Уитни». Авторы не проводят проверку ограничительных условий использования критерия Стьюдента. В тексте статьи присутствуют некорректные выражения с использованием слова «достоверно» [24].

39) ПКК, № 3, 2011, с. 67-70. Оценка состояния миокарда и ультраструктуры его микрососудов при хронической интоксикации опиатами и этанолом.

В разделе «Материал и методы» авторы пишут: «Материал был получен от 94 трупов лиц обоего пола, средний возраст погибших составил 30±1,5 года ($p>0,05$). … Контрольную группу составили 30 трупов лиц, умерших на догоспитальном этапе в результате механических повреждений (23 наблюдения) и механической асфиксии (7 наблюдений), их средний возраст составил 31±1,4 года ($p>0,05$)». Подобные обороты, содержащие среднее значение какого-то признака, и ошибку средней, а также выражения вида «$p>0,05$» встречаются в тексте статьи очень часто. Ниже приводим фрагмент текста из раздела «Результаты и обсуждение» с подобными оборотами. В разделе «Материал и методы» авторы сообщают: «При оценке структурно-функциональных изменений миокарда определяли средние величины морфометрических

показателей и ошибку средней (M ± m). Статистическую обработку количественных показателей проводили с помощью t-критерия Стьюдента, считая значимыми различия при p<0,05».



РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

При макроскопическом исследовании в основной группе масса сердца составила 380±17,4 г (p<0,05), (в группе контроля 330±21,9 г (p<0,05)). При осмотре полостей сердца отмечалось их умеренное расширение, так, ширина правого предсердия составила 3,7±0,1 см (p<0,05) (в группе контроля 2,4±0,1 см (p<0,05)) и высота 3,8±0,1 см (p<0,05) (в группе контроля 2,5±0,1 см (p<0,05)), ширина левого предсердия от 3,3±0,1 см (p<0,05), (в группе контроля от 2,6±0,1 см (p<0,05)) и высота 3,6±0,1 см (p<0,05) (в группе контроля 2,7±0,1 см (p<0,05)). Ширина правого желудочка в исследуемой группе составила 4,9±0,1 см (p<0,05) (в группе контроля 3,6±0,1 см (p<0,05)), высота 5,2±0,1 см (p<0,05) (в группе контроля 3,7±0,1 см (p<0,05)) и толщина 0,5±0,1 см (p<0,05) (в группе контроля 0,3±0,1 см (p<0,05)), ширина левого желудочка составила 5,9±0,2 см (p<0,05) (в группе контроля 4,7±0,2 см (p<0,05) и высота 5,8±0,1 см (p<0,05)

Таким образом, факт наличия выражения «p<0,05» означает, что авторы, используя t-критерий Стьюдента, проводили проверку неких статистических гипотез. Скорее всего, это была проверка гипотез о равенстве групповых средних. Однако абсолютно непонятно, для каких конкретно групп проводились эти сравнения, поскольку авторы утаили эту информацию, как от читателей, так и от редакции журнала. В результате подобные конструкции с выражениями «p<0,05» воспринимаются как «статистические гитики» [66]. В принципе возможны ситуации, когда t-критерий Стьюдента может использоваться и не для проверки статистических гипотез о равенстве средних двух групп. Его можно использовать и для проверки статистической гипотезы о равенстве генеральной (популяционной) средней некоторому конкретному числу, константе. Например, проверить гипотезу о том, что генеральное среднее равно нулю. И тогда вполне уместны будут такие выражения, как приведённые в статье, например, 380±17,4 г (p<0,05), и т.п.

Но разве имеет смысл гипотеза о равенстве массы сердца нулю, или о равенстве нулю его линейных размеров, и т.д.? О чём же говорит наличие подобных статистических гитик? С одной стороны, это показатель уровня знаний в области медицинской статистики самих авторов. Но с другой стороны, это ещё и показатель качества оценки содержания рукописей статьи, как редакцией самого журнала, так и рецензентами этой рукописи. Если, конечно, таковые имелись. Ведь ни у одного из тех, кто оценивал поступившую рукопись, не возникли подобные вопросы. Впрочем, вполне возможно, что как редакция журнала, так и рецензенты рукописи, вообще не считают себя вправе оценивать статистические детали поступающих рукописей, полностью оставляя эти аспекты на совести авторов.

40) ПКК, № 2, 2011, с. 17-20. Сравнительная оценка эластических свойств аорты у новорождённых с коарктацией аорты до и после хирургической коррекции.

В разделе «Материал и методы» авторы пишут: «Статистическая обработка полученных данных проводилась с помощью программы STATISTICA 6, при использовании t-теста для двух независимых выборок. Результаты представлены как среднее и стандартное отклонение (M ± σ). Статистически значимыми считались различия данных при p<0,001». Положительным моментом можно считать объяснение содержания выражений M ± σ. Однако авторы не провели проверку ограничивающих условий применения t-теста, в результате чего полученные с его помощью выводы нельзя считать надёжными. Далее, авторы никак не обосновывают выбор столь малого

критического уровня статистической значимости, равного 0,001. В статье используются некорректные выражения со словом «достоверно» [24].

41) ПКК, № 2, 2011, с. 55-58. Прекондиционирующий эффект севофлурана у больных ишемической болезнью сердца, оперированных в условиях искусственного кровообращения.

В разделе «Материал и методы» авторы пишут: «Статистический анализ проводили с помощью пакета программ Statistica 6.0. Для оценки характера распределения в совокупности по выборочным данным использовали тест Колмогорова – Смирнова. Сравнения двух групп из совокупностей с нормальным распределением проводили с помощью t-критерия Стьюдента для двух зависимых или двух независимых выборок. Для анализа зависимости количественных признаков применяли **ранговый коэффициент корреляции Пирсона**. Результаты представлены как среднее и стандартная ошибка среднего (M ± m). Статистически значимыми считались различия данных и корреляция при $p<0,05$». Положительным моментом описания можно считать достаточно подробное описание использованных методов статистики. Однако при этом авторы не проводили проверку второго ограничительного условия на использование критерия Стьюдента. Кроме того авторы не сообщают, с помощью каких методов и критериев проводилась проверка статистических гипотез в том случае, когда гипотеза нормальности отвергалась. Напомним, что одновременное выполнение обоих условий корректности использования критерия Стьюдента встречайте очень редко. При наличии же более двух групп сравнения, необходимо было учесть проблемы множественных сравнений, путём уменьшения критического уровня значимости, чего авторы не сделали. С учётом перечисленных выше причин результаты, полученные с помощью критерия Стьюдента, нельзя считать надёжными. Кроме того авторы путают два типа коэффициентов корреляции: коэффициент Пирсона и ранговый коэффициент Спирмена. Как и в предыдущей статье, авторы используют некорректные обороты с использованием слова «достоверно». [24]

42) ПКК, № 4, 2011, с. 23-28. Отдалённые результаты протезирования аортального клапана бескаркасными биопротезами «Кемерово-АБ-Моно», «Кемерово-АБ-Композит» и «Кемерово-АБ-Нео».

В разделе «Материал и методы» авторы пишут: «Применяли методы как параметрической, так и непараметрической статистики. Изучение проводили с учётом «Guidelines for Reporting Morbidity and Mortality After Cardiac Valvular Operations» [3]». Однако при этом авторы не конкретизируют, какие параметрические и непараметрические методы применялись ими, и для проверки каких конкретно статистических гипотез. Отсутствие этой информации не позволяет читателям оценить корректность использования авторами этих методов и критериев, а значит и надёжность полученных авторами выводов. В разделе «Материал и методы» авторы пишут: «При анализе отдалённых результатов основными методами были – метод множительных оценок Kaplan-Meier и метод таблиц и распределения времени жизни, а также **регрессионные модели**». Однако при этом авторы не уточняют, какие же конкретно регрессионные модели были ими использованы: модели Кокса, или же модели логистической регрессии [44]. Отметим, что для двух

этих регрессионных моделей структура результатов совершенно разная. Например, при использовании логистической регрессии, наиболее подходящей в данном случае, результат этого анализа содержит размерные коэффициенты регрессии и достигнутые уровни статистической значимости этих коэффициентов [44]. Кроме того, для полученного уравнения оценивается показатель согласия фактических и расчётных частот отнесения наблюдений к тем или иным исходам. Этот показатель называется коэффициентом конкордации. Чем ближе этот показатель к 1, тем адекватнее и точнее модель.

Помимо этих показателей данный метод позволяет оценить и так называемые безразмерные коэффициенты регрессии, ранжируя которые по их модулю, можно упорядочить предикторы (объясняющие переменные, переменные влияния) по силе их влияния на вероятность того или иного исхода наблюдения. К сожалению, ничего из перечисленного выше, авторы не приводят в тексте статьи. И по этой причине читателям невозможно установить, о каких конкретно регрессионных моделях авторы упоминают в разделе «Материал и методы». Единственное упоминание о результатах использования регрессионного анализа имеет следующий вид: «Для выявления факторов, определяющих величину ПЧГД на этих моделях в отдаленном периоде, мы провели многофакторный регрессионный анализ, в результате которого выявили, что наибольшее ($p<0,001$) влияние оказывает отношение диаметра протеза к диаметру фиброзного кольца аортального клапана. Зависимость эта имеет обратный характер – чем больше это отношение, тем меньше ПЧГД в отдаленном периоде». Из данного описания этого результата нельзя оценить следующие очень важные моменты. 1. Какова структура полученного уравнения (линейная или нелинейная зависимость). 2. Какова величина коэффициента корреляции между зависимой переменной ПЧГД и предиктором, и, соответственна величина коэффициента детерминации, равного квадрату коэффициента корреляции. Например, если коэффициент корреляции в данном случае будет равен 0,5 , то в этом случае коэффициент детерминации будет равен 0,5*0,5=0,25. Т.е. лишь 25% вариабельности ПЧГД будет определяться вариабельностью указанного признака, включённого в уравнение регрессии. Тогда как остальные 75% изменчивости ПЧГД будут определяться комбинацией значений всех остальных показателей, в т. ч. и показателей анамнеза, и других показателей. Вот как об этом пишут в близкой по тематике статье «Протезирование аортального клапана бескаркасными протезами «Кемерово-АБ-Композит Neo»: первый опыт» (А.М. Караськов, И.И. Семёнов, Д.А. Астапов, Е.И. Семёнова, В.Е. Железчиков, Д.В. Шматов. Кардиология и сердечно-сосудистая хирургия 2010; 5: 62-66). «Было установлено, что существует статистически значимая зависимость между величиной ПЧГД и методикой имплантации протеза ($p=0,02$), индексами КДО ЛЖ ($p=0,03$) и КДР ЛЖ ($p=0,004$) после операции. Однако фиксированный нелинейный регрессионный анализ с пошаговой моделью позволил выявить, что уровень статистической значимости регрессионных коэффициентов индексов КДР и КДО ЛЖ превышает принятое значение 0,05. Таким образом, наиболее значимым фактором, определяющим ПЧГД в раннем послеоперационном периоде, является методика имплантации бескаркасного биопротеза». Как видим, в данной статье авторы сообщают, что использовали нелинейный регрессионный анализ с пошаговым алгоритмом отбора предикторов.

Авторы анализируемой статьи «Отдалённые результаты протезирования …» также используют некорректное выражение «Это влияние статистически недостоверно» [24]. Далее авторы сообщают, «Средний срок наблюдения составил в общей группе 52 ± 29 (3–126) месяцев, … », однако при этом не уточняют, какая величина приводится ими после знака «±». Также не сообщается и то, какие величины приводятся ими в скобках. Вероятнее всего это минимальное и максимальное значение, а возможно и другие статистики, например, границы доверительного интервала, или квантили распределения, однако, так ли это в действительности, неизвестно. Если же вновь обратиться к статье «Протезирование аортального клапана бескаркасными протезами «Кемерово-АБ-Композит Neo»: первый опыт» (А.М. Караськов, И.И. Семёнов, Д.А. Астапов, Е.И. Семёнова, В.Е. Железчиков, Д.В. Шматов. Кардиология и сердечно-сосудистая хирургия 2010; 5: 62-66), то там под таблицей авторы пишут: «Примечание. * — данные представлены в виде среднего ± стандартное отклонение (минимум—максимум)». Можно предположить, что требования редакции журнала «Кардиология и сердечно-сосудистая хирургия» более чёткие и конкретные, нежели требования журнала ПКК, содержание которых мы обсудим далее. Либо различие в деталях описания статистических аспектов этих двух статей обусловлено различием авторских коллективов.

43) ПКК, № 4, 2011, с. 29-32. Тактико-технические аспекты комплексного лечения гнойного перикардита с синдромом сдавления сердца и полиорганной недостаточностью**.**

Авторы никак не описывают использованные методы статистического анализа собранных данных. При этом уже в аннотации данной публикации приводят выражения «средний возраст составил 42±16,6 лет » и «0,5 ± 0,1 балла (p=0,037)». Это говорит о том, что вероятнее всего авторы статьи использовали средние баллы, что является некорректно. При этом авторы не извещают читателей, что за величины размещены после знака ± .

44) ПКК, № 4, 2011, с. 33-38. Отдалённые результаты реконструктивных операций при постинфарктных аневризмах левого желудочка.

В разделе «Материал и методы» авторы пишут: «Статистическая обработка результатов проводилась на компьютере IBM Pentium 1400 с использованием пакета прикладных программ «Statistica 6.0» фирмы StatSoft Inc., t-критерия Стьюдента. Достоверными считались отклонения p<0,05. Данные расчётов представлены как M ± δ. Актуарная выживаемость определялась по методу Kaplan Meier». Упоминание марки компьютера «IBM Pentium 1400» не имеет никакого смысла для описания статистической обработки наблюдений. С таким же основанием помимо упоминания тактовой частоты в 1400 мегагерц, можно было перечислить и другие параметры компьютера. Например, число ядер центрального процессора (1, 2, 4, и т.д.) затем объём оперативной памяти, объём жесткого диска, название операционной системы, название материнской платы компьютера, марку винчестера (жёсткого диска), марку монитора, «мышки», и т.д. Использование t-критерия Стьюдента без проверки условий ограничения его применения автоматически делает полученные результаты сравнения не заслуживающими доверия. Не объясняют авторы и то, какие величины обозначены ими как M ± δ. Если величина «M» вероятнее всего есть среднее, то что такое «δ» – неизвестно. Это может быть полуширина

доверительного интервала (именно таким символом чаще всего в литературе обозначают полуширину доверительного (CI) интервала), либо выборочным стандартным отклонением, либо выборочной ошибкой среднего. Некорректно и использование авторами выражений вида «разница … статистически достоверна» [24].

45) ПКК, № 4, 2011, с. 53-56. Сравнительная характеристика поражения коронарных артерий у больных ишемической болезнью сердца с метаболическими факторами риска при наличии либо отсутствии абдоминального ожирения.

Авторы приводят выражения вида «M ± m = 32,2±0,4», однако не конкретизируют что за величины связаны знаком « ± » . В статье часто используются выражения вида «P < 0,05», из чего можно сделать вывод о том, что авторы проводили проверку неких статистических гипотез. Однако при этом авторы скрывают от читателей, как сами проверяемые гипотезы, так и статистический инструментарий их проверки (статистические критерии, проверка условий их использования, критические уровни значимости и т.п.). Не сочла необходимой эту информацию и редакция журнала ПКК, принявшая решение о публикации данной статьи. В статье используется некорректные выражения вида «статистически достоверного различия не наблюдалось» [24]. Следует отметить, что выражения вида «P < 0,05» используются как в данной статье, так и в большинстве проанализированных выше статей. И это при том, что авторы таких статей отнюдь не рядовые врачи отделения клиник, диспансеров, или городских больниц. В большинстве случаев это профессиональные исследователи, имеющие кандидатские или докторские степени, профессора, преподаватели медицинских вузов, директора НИИ, или их заместители по науке, и т.д. То есть это специалисты, имеющие непосредственное отношение, как к науке, так и к преподаванию в вузах. А некоторые из них даже ведут курсы основ доказательной медицины (ДМ). В частности, первый из авторов данной статьи является руководителем курса «Основы доказательной медицины» кафедры **внутренних болезней стоматологического факультета** Новосибирского государственного медицинского университета. В докладе на Ереванской конференции по ДМ [46] мы также обсуждали другую статью этих авторов. Из которой привели несколько примеров некорректного использования статистики.

Обратимся к наиболее известным и доступным изданиям по основам доказательной медицины, таким как «Клиническая эпидемиология. Основы доказательной медицины», авторы Р. Флетчер, С. Флетчер, Э. Вагнер (М.: Медиа Сфера, 1998. – 352 с.), и «Введение в доказательную медицину», автор В.В. Власов (М.: Медиа Сфера, 2001. – 392 с.). Уже от знакомства только с двумя этими книгами становится понятно, что необходимо в публикациях конкретно формулировать какие статистические гипотезы проверяются, и что необходимо не просто записывать P < 0,05, а писать фактические, достигнутые значения уровня значимости указываемого статистического критерия. Вот как пишет об этом в своей книге Президент Общества специалистов доказательной медицины, д.м.н., профессор В.В. Власов (стр. 121–123). «Чаще всего исследователя интересует различие, и соответственно речь идёт об отклонении нулевой гипотезы. Поэтому обычно оценивают именно риск принятия ошибочного решения о том, что различие существует (ошибка первого рода, или α–ошибка; табл. 4.1). Этот риск называют α–риском. Для него-то и устанавливается пороговая величина вероятности ошибки *p* на уровне,

традиционно равном 0,05 или 0,01. Публикации легче оценивать, если в них вероятность ошибки *p* приводится не в отношении значения ($p< 0,05$), а в виде точной величины ($p=0,0017$). При использовании современных вычислительных средств точное вычисление *p* не составляет проблемы; большинство статистических программ позволяют рассчитывать величину *p* . При $p<0,0001$ точная величина уже не имеет значения, поскольку для медицинских решений это слишком низкие вероятности. Если вероятность ошибки первого рода близка к пороговой, то при обсуждении различий не может быть разного подхода к **случаям**, когда $p=0,04$ и $p=0,06$. Принцип оценки таких результатов должен быть сходным, несмотря на то, что формально один достигает условного уровня статистической значимости, а другой – нет».

Итак, когда медики, причисляющие себя к сторонникам ДМ, и ведущие курсы «Основы доказательной медицины», допускают столь грубое игнорирование принципов ДМ, то по этим фактам можно судить как об уровне знания ими принципов ДМ, так и об уровне их знаний в статистике.

46) ПКК, № 1, 2012, с. 11-15. Динамика процессов адаптации левого желудочка у больных раннего возраста с коарктацией аорты до и после хирургической коррекции.

«Статистическая обработка полученных данных проводилась с помощью программы Statistica 6.0 с использованием t-теста для двух независимых выборок. Результаты представлены как среднее и стандартное отклонение (M ± σ). Статистически значимыми считались различия данных при $p<0,05$». Положительным моментом этого описания можно считать конкретизацию величин в выражениях вида M ± σ. Однако выборочное стандартного отклонение следует обозначать не греческой буквой «σ» (сигма), а латинской буквой «s». Поскольку авторы не проводили проверку условий корректного применения t-теста, то выводы, полученные с его помощью, нельзя считать заслуживающими доверия. Учитывая наличие набора различных ЭхоЛГ ЛЖ, а также нескольких групп сравнения, следовало не ограничиваться лишь примитивным сравнением средних в этих группах, а использовать набор многомерных методов статистического анализа.

47) ПКК, № 1, 2012, с. 23-26. Протезирование аортального клапана бескаркасным биопротезом «Кемерово-АБ-Нео»: непосредственные результаты.

«Для оценки типа распределения данных использовали параметр Шапиро – Вилка. Применяли методы параметрической статистики. **Уровень достоверности** принимали равным 0,05. При многофакторном анализе взаимосвязей сначала путем однофакторного изучения выделяли основные параметры, влияющие на исследуемую величину, затем на основании поиска межгрупповых корреляций отсеивали признаки, имеющие умеренную или сильную связь между собой, и проводили многофакторное моделирование взаимосвязей». Авторы сообщают о том, что использовали методы параметрической статистики. Но при этом умалчивают, какие конкретно методы были ими использованы. Учитывая тот факт, что для проверки гипотезы о типах распределения авторы использовали метод Шапиро-Вилка, утверждение об использовании параметрических методах статистики автоматически означает, что все проверки подтвердили ожидаемый закон распределения. Можно предположить, что авторы подразумевали проверку нормальности распределения. Однако одной лишь нормальности распределения для многих

методов недостаточно. Например, для использования t-критерия Стьюдента и параметрического дисперсионного анализа (ANOVA) необходимо выполнение и второго ограничивающего условия. О чём авторы не сообщают, на основании чего можно сделать вывод о том, что это второе условие не проверялось.

Не сообщают авторы также и о том, с помощью каких конкретно статистических методов они проводили многофакторный анализ взаимосвязей. Непонятно также и выражение о поиске межгрупповых корреляций. В частности, каким методом оценивались корреляции между группами, и между какими группами. Например, в биостатистике есть метод канонической корреляции. В этом методе оценивается корреляция между двумя группами количественных переменных. В таблице авторы приводят 4 группы анализируемых показателей. Из них 3 группы имеют значения по всем 4 признакам. В этом случае имеем 3 пары групп признаков, для которых может быть оценена каноническая корреляция. Это пары 1-2, 1-3 и 2-3. Однако по имеющей в статье информации невозможно предположить, о каких группах пишут авторы: о группах наблюдений, или же о группах признаков.

Таким образом, использованные авторами статьи туманное и расплывчатое описание статистических аспектов вызывает массу вопросов. Разумеется, такие вопросы возникают лишь у тех читателей, и у членов редколлегии ПКК, которые имеют хотя бы некоторое представление о возможностях статистики. Далее в статье авторы приводят ещё одно описание использованного ими статистического метода. «Многомерный нелинейный регрессионный анализ взаимосвязей с пошаговой моделью и обязательным контролем межвариантных корреляций показал, что уровень значимости регрессионных коэффициентов индексов КДР и КДО ЛЖ больше 0,05». И в этом описании тоже неясно, какая конкретно регрессионная модель оценивалась авторами. Следует напомнить, что количество нелинейных моделей в принципе в сотни и тысячи раз больше, чем количество линейных моделей. И уже в силу того, что нелинейные модели сложнее линейных, необходимо приводить более подробную информацию как о структуре модели, хотя бы в самом общем виде, так и давать описание использованных алгоритмов. Например, авторы сообщают об использовании пошаговой модели. Между тем пошаговый алгоритм оценки уравнения регрессии не один, а их несколько. И все они могут давать различные уравнения. Именно поэтому в целях получения оптимального варианта уравнения и следует использовать все эти алгоритмы.

О чём же свидетельствуют столь краткие описания сложных использованных методов? В статье «КРАТКОСТЬ – СЕСТРА ТАЛАНТА? ИЛИ ПРИЗНАК НЕЗНАНИЯ?» (http://www.biometrica.tomsk.ru/kuzbass3.htm) , мы называли основные причины такой краткости. Чаще всего это просто недостаточный уровень знаний авторов об используемых ими методах, либо опасение допустить многочисленные ошибки в более подробном описании использованных методов статистического анализа.

48) ПКК, № 1, 2012, с. 33-37. Эндотелиальная дисфункция и фактор курения у пациентов с ишемической болезнью сердца при тестировании лучевой артерии перед коронарным шунтированием.

«Статистическая обработка полученных результатов проводилась с использованием программного пакета Statistica 6.1. Для определения статистической значимости различий применялись **методы вариационной статистики**. Данные представлены в виде средних значений ± ошибка среднего (M ± m). Различия считали достоверными при уровне значимости p<0,05». Положительным моментом данного описания является конкретизация выражения M ± m, а также указание критического уровня статистической значимости.

Однако самая грубейшая ошибка данной публикации заключается в отсутствии перечня конкретных статистических методов использованных авторами. Выражение «Для определения статистической значимости различий применялись **методы вариационной статистики**» не содержит никакой полезной информации. Действительно, о различии чего, каких параметров, идёт речь? Неизвестно. Далее, бывает ли статистика без вариации? Процитируем фрагмент нашей статьи [38]. «В таких работах вместо конкретных наименований статистических методов и критериев авторы упоминают лишь одно словосочетание "**вариационная статистика**", введенное в отечественный оборот Ю.А. Филипченко ещё в 1918 г. [14]. Этот термин возник как неудачный, дословный перевод немецкого слова "Variationsstatistik", которое разумнее было бы перевести как "статистика вариаций" или "статистика изменчивости". Однако уже в тот период ряд авторов использовали словосочетание "биологическая статистика". Закреплению этого неудачного термина способствовало и издание в 1961 г. книги академика Рокицкого П.Ф. "Основы вариационной статистики для биологов". Однако **уже в следующем издании этой книги вышедшей в 1973 г. под названием "Биологическая статистика"**, автор пишет: "**Мы сочли целесообразным отказаться от термина "вариационная статистика"**. В свое время он был очень распространён, но сейчас употребляется довольно редко, т.к. содержит элементы тавтологии (статистический метод обязательно предусматривает и изучение вариации). С другой стороны, применение статистических методов в биологии приобрело такие особенности, что **можно с полным правом говорить о биологической статистике как самостоятельной области статистики**. Это и явилось основанием назвать книгу «Биологическая статистика» [40]». Таким образом, ещё 40 лет тому назад известный учёный, автор самой популярной книги по статистике в биологии и медицине, академик Рокицкий П.Ф. отказался от неудачного термина «вариационная статистика», и стал использовать более приемлемый термин «биологическая статистика». Из чего можно сделать вывод о том, что авторы, использующие выражение «вариационная статистика» вообще не подозревают о существовании нового, современного термина. И это же свидетельствует о незнании ими существования книги Рокицкого П.Ф. «Биологическая статистика».

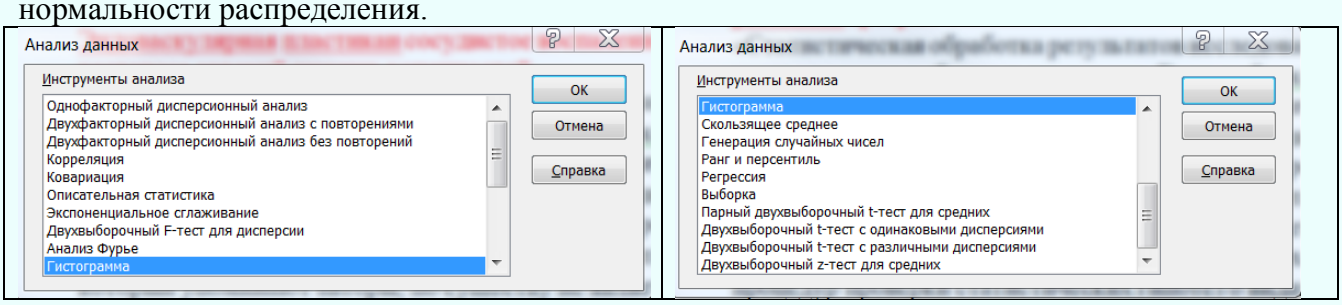
49) ПКК, № 1, 2012, с. 39-42. Психогенная одышка и гипокапния у больных ишемической болезнью сердца до и после коронарного шунтирования

«**По возрасту больные и здоровые не различались (p>1).** Статистическая обработка полученных данных проведена с использованием метода вариационной статистики. Определялись среднее арифметическое значение M, средняя ошибка арифметического значения (m) и t-критерий Стьюдента. Различия считались достоверными при p<0,05. Математические расчеты и вычисления статистических показателей проводились с использованием пакета прикладных программ Statistica 6.0». Судя по контексту в выражении **p > 1** символ «p» обозначает вероятности, т.е. величину

достигнутого уровня статистической значимости при проверке статистической гипотезы о равенстве средних возрастов в группе больных и в группе здоровых. Из чего следует, что авторы статьи не знают, что в современной теории вероятностей величина «p» может принимать значения только в интервале от 0 до 1. Т.е. не существует вероятности более 1. И в этом случае выражение **p > 1** просто бессмысленно. Эта статистическая нелепость вызывает большое недоумение ещё и потому, что среди авторов статьи 2 доктора наук. Помимо 2-х докторов наук в авторах статьи также и с.н.с. **Института вычислительной техники** СО РАН. Не могут вызывать доверия у читателей статьи и результаты, полученные авторами с помощью t-критерия Стьюдента, поскольку авторы не проводили проверку условий корректности использования данного критерия. Итак, что же это: ошибка авторов, или же это ошибка редакции журнала ПКК?

50) ПКК, № 1, 2012, с. 51-55. Эндоваскулярная пластика и сосудистое воспаление: взаимосвязь и влияние на развитие рестеноза артерий нижних конечностей.

«Статистическая обработка результатов исследования выполнялась с учётом вида **распределения выборок и дисперсии. Если выборка отличалась от нормальной,** использовались непараметрические методы. Различия считались достоверными при p<0,05. Вычисление результатов выполнялось при помощи пакета статистических программ Excel». Каков смысл выражения «с учётом вида распределения выборок и дисперсии»? Это полнейшая бессмыслица, которая подтверждается также и тем, что пакет программ Excel не является «пакетом статистических программ», хотя и содержит небольшой набор статистических функций. Однако, среди этих функций нет процедур проверки статистических гипотез о видах (законах) распределения вероятностей. Ниже приведён состав статистических процедур из надстройки «Анализ данных» пакета Excel. Как видим, нет там процедуры проверки гипотезы, например, о нормальности распределения.



Нет этих функций, а также и непараметрических методов, и среди всего набора 98 статистических функций пакета Excel. Ниже в таблице приведены 2 фрагмента (начальный и конечный списки) этого перечня статистических функций.

	A	B
1	Функция F.ОБР	Возвращает обратное значение для F-распределения
2	Функция F.ОБР.ПХ	Возвращает обратное значение для F-распределения
3	Функция F.РАСП	Возвращает F-распределение вероятности.
4	Функция F.РАСП.ПХ	Возвращает F-распределение вероятности.
5	Функция F.ТЕСТ	Возвращает результат F-теста.

94	Функция ХИ2.РАСП.ПХ	Возвращает одностороннюю вероятность распределения хи-
95	Функция ХИ2.ТЕСТ	Возвращает тест на независимость.
96	Функция ЧАСТОТА	Возвращает распределение частот в виде вертикального
97	Функция ЭКСП.РАСП	Возвращает экспоненциальное распределение.
98	Функция ЭКСЦЕСС	Возвращает эксцесс множества данных.

Возникает вопрос, зачем же авторам понадобилось убеждать читателей в необходимости изучения ими «вида» (закона) распределения дисперсии? Ведь уже более 100 лет как известно, что дисперсии распределены по распределению Пирсона Хи-квадрат. Критерий Хи-квадрат был предложен Карлом Пирсоном (Pearson) ещё в 1900 году. Видимо авторы статьи, как и редакция журнала ПКК, об этом не знают. Другая некорректность обсуждаемого выражения заключается в том, что при проверке гипотез о законе распределения, данные гипотезы формулируются не о

законе распределения выборки, а о законе распределения в генеральной совокупности (популяции), откуда была извлечена случайная выборка. И если гипотеза нормальности принимается, то это означает, что именно для генеральной совокупности (популяции), а не для выборки, не отвергается гипотеза нормальности.

Таким образом, смутное описание авторами выполненных ими статистических процедур не способно вызвать у вдумчивого читателя доверия к авторским результатам. В этой связи возникает вопрос, какова же была цель этой публикации, и на каких читателей ориентировали эту публикацию авторы? Не исключено, что подобные проблемы обусловлены авторским составом данной статьи, в котором на первом месте стоит автор, не имеющий учёной степени, а два доктора медицинских наук расположены ниже. В том, что такое расположение соавторов не определяется азбукой, можно судить по последовательности практически всех статей журнала ПКК. Тем более, что такой порядок расположения соавторов не предусмотрен редакционными требованиями. Если сделать несложный поиск в сети интернет, то мы узнаем, что в мае 2012 г. первый соавтор этой статьи защитил в НИИ им. академика Е.Н. Мешалкина кандидатскую диссертацию по теме «ЧРЕСКОЖНАЯ АНГИОПЛАСТИКА, ПРОВОСПАЛИТЕЛЬНЫЕ ФАКТОРЫ И ЭНДОТЕЛИАЛЬНАЯ ДИСФУНКЦИЯ В ЭФФЕКТИВНОСТИ РЕВАСКУЛЯРИЗАЦИИ НИЖНИХ КОНЕЧНОСТЕЙ».

Этот же поиск приводит и к статье «Воспаление сосудистой стенки и рестеноз после ангиопластики артерий нижних конечностей», опубликованной в журнале «Медицина и образование в Сибири», 2011, № 6.(http://www.ngmu.ru/cozo/mos/article/text_full.php?id=560) В этой статье читаем: «Статистическую обработку материала производили по **стандатной** (так написано в оригинале статьи. В.Л.) программе с вычислением достоверности различий по t-критерию Стьюдента. Связь между исследуемыми параметрами оценивали с помощью коэффициентов корреляции®. **Статистически значимыми считали различия более 0,05**». Напомним, что в статье «Эндоваскулярная пластика и сосудистое воспаление: взаимосвязь и влияние на развитие рестеноза артерий нижних конечностей» (ПКК, № 1, 2012, с. 51-55) эти же самые авторы утверждали: «**Различия считались достоверными при p<0,05**». Не имеет смысла пытаться понять, почему это вдруг 3 автора, 2 из которых доктора медицинских наук, вдруг изменили свою точку зрения в данном вопросе на противоположную. И какой стандарт они подразумевали в выражении «Статистическую обработку материала производили по **стандатной** программе».

Об оригинальности же каждой из двух статей этих авторов, читатели могут судить по следующим фрагментам. Вначале приведём таблицу №1 из статьи «Воспаление сосудистой стенки и рестеноз после ангиопластики артерий нижних конечностей», опубликованной в журнале «Медицина и образование в Сибири», 2011, № 6. А ниже приведём таблицу № 1 из статьи «Эндоваскулярная пластика и сосудистое воспаление: взаимосвязь и влияние на развитие рестеноза артерий нижних конечностей» (ПКК, № 1, 2012, с. 51-55).

Таблица 1

Концентрация провоспалительных факторов (вчСРБ, ИЛ-6, ЭТ-1, ФВ) в сыворотке крови больных атеросклерозом нижних конечностей (M ± m)

Группы обследованных пациентов	Количество пациентов (n)	Исследуемые показатели			
		вчСРБ (мг/л)	ИЛ-6 (пг/мл)	ЭТ-1 (нг/мл)	ФВ (%)
Больные АНК	128	11,7 ± 3,52a	7,92 ± 1,06a	1,24 ± 0,31a	115,7 ± 18,64
Контрольная группа	48	3,5 ± 0,37	1,85 ± 0,19	0,26 ± 0,03	84,8 ± 21,43

Прочитай и перешли своим коллегам…

Таблица 1
*Концентрация маркеров воспаления и повреждения эндотелия, * p<0,01 по сравнению с контролем, ** p<0,02 по сравнению с данными до операции*

Группы	Показатели			
	вчСРБ, мг/л	ИЛ-6, пг/мл	ЭТ-1, нг/мл	ФВ, %
Больные АНК, n = 128				
до операции	11,7±3,52*	7,92±1,06*	1,24±0,31*	115,7±18,64
3 сут. после операции	21,2±1,72**	13,1±1,65**	2,4±0,37**	146,2±15,23
Контрольная группа, n = 48	3,5±0,37	1,85±0,19	0,26±0,03	84,8±21,43

Далее приведём таблицу из статьи «Воспаление сосудистой стенки и рестеноз после ангиопластики артерий нижних конечностей», опубликованной в журнале «Медицина и образование в Сибири», 2011, № 6.

А ниже приводим аналогичную таблицу из статьи «Эндоваскулярная пластика и сосудистое воспаление: взаимосвязь и влияние на развитие рестеноза артерий нижних конечностей» (ПКК, № 1, 2012, с. 51-55).

Концентрация провоспалительных факторов (вчСРБ, ИЛ-6, ЭТ-1, ФВ) в сыворотке крови больных АНК в первые 3 суток, а также спустя 3, 6 и 12 месяцев после операции (M ± m)

Группы обследованных пациентов	Кол-во обследованных n, (%)	Период исследования	Исследуемые показатели			
			вчСРБ (мг/л)	ИЛ-6 (пг/мл)	ЭТ-1 (нг/мл)	ФВ (%)
Пациенты после ангиопластики подвздошных артерий без рестеноза	49 (38)	3 месяца после операции	16,5 ± 1,55	10,1 ± 0,35	2,1 ± 0,21	132,1 ± 15,32
		6 месяцев после операции	16,1 ± 2,52	9,9 ± 0,47	1,8 ± 0,15	129,2 ± 16,21
		12 месяцев после операции	15,2 ± 2,35a	9,6 ± 0,58a	1,5 ± 0,25a	123,4 ± 17,52
Пациенты после ангиопластики подвздошных артерий с наличием рестеноза	25 (19,5)	3 месяца после операции	19,1 ± 1,27	11,9 ± 1,14	2,3 ± 0,31	141,5 ± 16,25
		6 месяцев после операции	18,9 ± 1,73	11,1 ± 1,32	2,1 ± 0,27	138,1 ± 14,11
		12 месяцев после операции	18,3 ± 1,95	10,7 ± 1,24	1,9 ± 0,43	132,7 ± 15,21
Пациенты после ангиопластики бедренных и подколенных артерий без рестеноза	29 (23)	3 месяца после операции	17,9 ± 1,25	9,9 ± 0,59	2,0 ± 0,17	131,2 ± 16,71
		6 месяцев после операции	17,2 ± 2,17	9,5 ± 0,72	1,8 ± 0,12	129,5 ± 17,23
		12 месяцев после операции	16,1 ± 1,88a	9,1 ± 0,93a	1,6 ± 0,15a	125,6 ± 15,31
Пациенты после ангиопластики бедренных и подколенных артерий с наличием рестеноза	25 (19,5)	3 месяца после операции	20,9 ± 1,44	12,8 ± 0,52	2,3 ± 0,21	142,3 ± 15,31
		6 месяцев после операции	20,2 ± 1,32	12,1 ± 0,64	2,2 ± 0,25	138,1 ± 16,72
		12 месяцев после операции	19,1 ± 1,23	11,2 ± 0,79	2,1 ± 0,23	132,7 ± 15,21
Данные раннего послеоперационного периода (контроль)	128 (100)	Первые трое суток после операции	21,2 ± 1,72	13,1 ± 1,65	2,4 ± 0,37	146,2 ± 15,23

Таблица 2
*Концентрация маркеров воспаления и повреждения эндотелия, * p<0,01 по сравнению с контролем*

Пациенты	После операции, мес.	Показатели вчСРБ, мг/л	ИЛ-6, пг/мл	ЭТ-1, нг/мл	ФВ, %
После ангиопластики					
подвздошных артерий без рестеноза, n = 49	3	16,5±1,55	10,1±0,35	2,1±0,21	132,1±15,32
	6	16,1±2,52	9,9±0,47	1,8±0,15	129,2±16,21
	12	15,2±2,35*	9,6±0,58*	1,5±0,25*	123,4±17,52
подвздошных артерий с рестенозом, n = 25	3	19,1±1,27	11,9±1,14	2,3±0,31	141,5±16,25
	6	18,9±1,73	11,1±1,32	2,1±0,27	138,1±14,11
	12	18,3±1,95	10,7±1,24	1,9±0,43	132,7±15,21
бедренных и подколенных артерий без рестеноза, n = 29	3	17,9±1,25	9,9±0,59	2,0±0,17	131,2±16,71
	6	17,2±2,17	9,5±0,72	1,8±0,12	129,5±17,23
	12	16,1±1,88*	9,1±0,93*	1,6±0,15*	125,6±15,31
бедренных и подколенных артерий с рестенозом, n = 25	3	20,9±1,44	12,8±0,52	2,3±0,21	142,3±15,31
	6	20,2±1,32	12,1±0,64	2,2±0,25	138,1±16,72
	12	19,1±1,23	11,2±0,79	2,1±0,23	132,7±15,21
Ранний послеоперационный период (контроль), n = 128 (100%)	3 суток	21,2±1,72	13,1±1,65	2,4±0,37	146,2±15,23

51) ПКК, № 3, 2012, с. 13-17. Химические элементы и структурно-молекулярные особенности кардиомиоцитов у пациентов раннего возраста с транспозицией магистральных артерий.

«Статистическую обработку результатов проводили с помощью программы Microsoft Excel 2000. Достоверность различий средних величин и корреляционных взаимоотношений проводили с помощью t-критерия Стьюдента. Достоверными считали различия при p<0,005». Некорректности данного описания включают использование t-критерия Стьюдента без проверки условий его использования. Неясно, какие конкретно коэффициенты корреляции использовались авторами. Следует отдельно остановиться на особенности тематики данного исследования. Тема исследования бесспорно очень важна и интересна не только с чисто теоретической стороны, но и с практической. Несложно понять, что трудоёмкость формирования массива данных для проведения его анализа весьма велика. Поскольку требует использования очень тщательных и дорогих методик. И уже в силу этого анализ столь ценного массива данных требует не менее трудоёмкого и глубокого статистического анализа. К большому сожалению, авторы исследования ограничились лишь сравнением средних и оценкой корреляции. Ниже приведён фрагмент статьи, в котором авторы приводят значения корреляций. Однако, не уточняя при этом, что это за коэффициенты.

Таблица 4
Корреляция между толщиной миокарда и содержанием в нем некоторых химических элементов
** p<0,05; ** p<0,01*

Желудочек	Анатомические группы	S	K	Fe	Cu	Sr	Zn
Правый	I	0,15	0,09	0,58*	-0,26	-0,19	0,12
	II	0,31	0,15	-0,29	-0,60*	0,13	0,08
Левый	I	0,92**	0,75**	0,82**	-0,92**	0,67**	-0,18
	II	0,46	0,47	0,50	-0,33	0,30	-0,25

Напомним, что существует порядка десяти различных коэффициентов корреляции. В том числе, для линейных и нелинейных парных зависимостей. Далее, уже один лишь факт наличия статистически значимых корреляций свидетельствует о том, что корреляционные связи несут в себе немалую толику полезной информации. И по этой причине целесообразно помимо сравнения средних, использовать и многомерные методы анализа. Например, в данной ситуации целесообразно использовать такие методы, как факторный анализ, дискриминантный анализ, кластерный анализ, логистическую регрессию,[44] каноническую корреляцию, и т.д. Увы, судя по всему, авторы статьи не владеют этими методами, и поэтому большая часть ценнейшей информации так и осталась не извлечённой из собранного массива данных.

52) ПКК, № 3, 2012, с. 53-56. Использование стратегии преднамеренного многоэтапного комбинированного лечения у пациентов со множественными церебральными аневризмами в остром периоде кровоизлияния.

Прочитай и перешли своим коллегам…

«Статистический анализ проводился с использованием пакета прикладных программ Statistica v.6.0». Однако при чтении текста статьи там не обнаруживается никаких признаков проведения статистического анализа. В тексте приведены 25 раз процентные соотношения, для вычисления которых вовсе не требуется статистический пакет программы Statistica v.6.0. Поскольку вычисление процентов можно проводить с помощью простого калькулятора.

53) ПКК, № 3, 2012, с. 67-74. Васкулярно-паренхиматозные взаимоотношения в правом предсердии при различных способах гипотермии при операциях на открытом сердце в клинике и эксперименте.

«Для анализа ультраструктуры митохондрий (МХ) каждый их хорошо видимый профиль оценивали **по трехбалльной системе**…. **Результаты представляли в виде среднего значения и его ошибки**. … Статистическую обработку полученных результатов выполняли с применением **параметрических методов анализа.** Во всех группах достоверность изменений ультраструктуры МС и КМЦ на этапах кардиохирургических вмешательств определяли **с помощью t-критерия Стьюдента**. Вычисление коэффициента корреляции осуществляли по формулам, рекомендованным медико-биологической статистикой для оценки взаимосвязи между малочисленными группами при малозначных датах [8]». Авторы статьи не провели проверку корректности использования критерия Стьюдента. Между тем одновременное выполнение 2 ограничений использования этого критерия встречается в реальной практике весьма редко, примерно в 2-5% случаев. Обратим также внимание на то, что авторы использовали t-критерия Стьюдента по отношению к признакам, которые в принципе не могут нормального распределения. Поскольку анализируемые признаки «**оценивали по трехбалльной системе**», Т.е. признаки не являлись непрерывными. И, стало быть, для балльных признаков в принципе нельзя применять арифметические операции сложения, вычитания, умножения и деления, которые обязательны при вычислении критерия Стьюдента. Т.е. авторам априорно нужно было использовать непараметрические статистические критерии. Из чего можно сделать заключение об отсутствии доверия к результатам сравнения средних с помощью этого критерия. Не уточняют авторы и то, какой конкретно вид коэффициента корреляции был ими использован. Учитывая то, что в исследовании использовались балльные показатели, необходимо было использовать ранговый коэффициент корреляции Спирмена. Однако авторы не уточняют, какой именно вид коэффициента корреляции был ими использован.

54) ПКК, № 4, 2012, с. 21-26. Многолетний опыт хирургии сонных артерий для профилактики ишемического инсульта.

«Статистическая и аналитическая обработка полученных материалов, стандартизация исследования проводились **по критериям методического руководства Сибирского межведомственного регионального Центра по регистру и профилактике инсульта** (ННИИПК им. акад. Е.Н. Мешалкина Росздрава и Института вычислительной математики и математической геофизики СО РАН, 2001–2003)». Это краткое описание использованных статистических методов обработки материалов нельзя признать удовлетворительным, потому как авторы не сообщают ни выходных данных этого методического руководства, ни названия конкретных использованных методов анализов. Авторы сообщают, «Средний возраст в группах хирургического лечения составил 55±7,1 года и **статистически достоверно** не различался в подгруппах исследования ($p>0,05$), по данным дисперсионного анализа ANOVA». Однако при этом ничего не сообщают о проверке двух условий корректности использования метода ANOVA. В силу чего полученные ими с помощью этого методов выводы можно априори считать сомнительными. Поскольку

одновременное выполнение 2-х условий корректности в реальной практике встречается крайне редко. Авторы статьи используют выражения вида «59,5±7,7», однако при этом не уточняют, что за величины представлены в таких выражениях.

55) ПКК, № 4, 2012, с. 43-47. Применение сандостатина в профилактике острого воспалительного ответа при операциях реваскуляризации миокарда в условиях искусственного кровообращения.

«Статистический анализ полученных результатов провед`н с помощью программы Statistica 6,0. Для оценки характера распределения в совокупности по выборочным данным использовали тест Колмогорова – Смирнова. Сравнение групп из совокупностей с нормальным распределением проводили с помощью t-критерия Стьюдента для двух зависимых или независимых выборок. Результаты представлены как средние и стандартные ошибки средних (M ± m). Статистически значимыми считали различия данных при $p<0,05$». Согласно этому описанию выходит, что все признаки во всех сравниваемых подгруппах имели нормальное распределение, что весьма маловероятно. Кроме того, авторы не сообщают о проверке второго обязательного условия для корректного применения t-критерия Стьюдента. Исходя из этого, следует считать выводы, полученные с помощью критерия Стьюдента, нельзя считать надёжными. Положительным элементом публикации является конкретизация выражения M ± m, а также использование критерия Колмогорова-Смирнова.

56) ПКК, № 1, 2013, с. 11-14. Оценка ремоделирования полостей сердца у пациентов с острой тромбоэмболией легочной артерии методом трансторакальной эхокардиографии.

«Статистическая обработка материала проводилась с помощью программы Statistica 6. Результаты представлены как среднее и среднеквадратичное отклонение (M ± σ). Достоверность различий средних величин определяли по парному и непарному критерию Стьюдента (t). Статистически значимыми считались различия данных при $p<0,05$». Авторы не проверяли выполнение условий корректности использования t-критерия Стьюдента. Кроме того, в тексте статьи авторы приводят таблицы с тремя группами сравнения. Из чего следует, что в этом случае необходимо учитывать проблему множественных сравнений, уменьшая для этого величину критической значимости. Однако и этого авторы не делают. Из чего можно сделать заключение о сомнительности выводов, полученных авторами с помощью критерия Стьюдента. Некорректно также и использование авторами выражений типа «достоверные различия» [24].

57) ПКК, № 1, 2013, с. 15-20. Сравнительная характеристика эффективности методов непрямой реваскуляризации миокарда в хирургии ишемической болезни сердца.

«Полученные результаты обрабатывались с помощью программы «STATISTICA 6». Учитывая нормальное распределение выборок, использовали Т-критерии для независимых и зависимых выборок, коэффициент корреляции Пирсона. Данные представлены в виде M ± m (среднее значение ± стандартная ошибка). Значение $p\leq0,05$ считали статистически достоверным». Положительным моментом описания следует считать конкретизацию выражений M ± m, конкретизацию использованного коэффициента корреляции, и указание критического уровня статистической значимости, равного 5%. Однако авторы не сообщают, каким образом они проверяли условия корректного использования критерия Стьюдента. Которые, как известно, заключаются в нормальности распределения в обеих группах сравнения, а также в равенстве генеральных дисперсий. Непонятно, как «учитывалось нормальное распределение выборок»?

Прочитай и перешли своим коллегам…

Напомним также, что проверка гипотез о нормальности закона распределения вероятностей делается С ИСПОЛЬЗОВАНИЕМ ВЫБОРОК, но при этом сама статистическая гипотеза нормальности формулируется НЕ ОТНОСИТЕЛЬНО ВЫБОРКИ а относительно ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ (ПОПУЛЯЦИИ) из которой извлекались эти выборки. Кроме того, из таблиц, приведённых в тексте статьи, следует, что в анализе использовались 5 групп сравнения: «до операции», «через 2 недели после операции», «через 6 месяцев после операции», «через 12 месяцев после операции», и «через 3 года после операции». В этом случае необходимо учитывать проблему множественных сравнений, для чего следует уменьшать критический уровень значимости. Чего авторы не делают. В силу чего декларируемые авторами выводы нельзя признать надёжными.

58) ПКК, № 1, 2013, с. 61-65. Результаты одномоментных операций протезирования восходящего отдела аорты и реваскуляризации миокарда.

В разделе «Материал и методы» авторы публикации ничего не сообщают об использованных ими статистических методах. Все результаты анализа данных по 264 пациентам представлены в виде 30 выражений типа 56,8 ± 5,8 и более 100 процентных отношений. Есть также и выражения типа «r = 0,497; p = 0,002». Можно предположить, что в данном случае речь идёт о величине коэффициента корреляции. Однако при этом авторы не уточняют, какой конкретно коэффициент корреляции оценивался ими.

59) БЮЛЛЕТЕНЬ СО РАМН, №2 (120), 2006 г. Хирургическое лечение мультифокального атеросклероза.

«Выявлены **достоверные различия** частоты встречаемости осложнений после каротидной ангиографии и частоты осложнений после компьютерной ангиографии ($p<0,05$) по данным дисперсионного анализа для малых выборок (при использовании критерия χ^2 и критерия χ^2 с поправкой Йетса) [10]». Авторы утверждают, что для оценки различия частот в двух группах сравнения (после каротидной ангиографии и после компьютерной ангиографии) они использовали дисперсионный анализ. Однако при этом не уточняют, какие конкретно виды дисперсионного анализа были ими использованы. Ведь помимо классического дисперсионного анализа (ANOVA) существуют и иные виды непараметрического дисперсионного анализа. Также непонятно, почему вместо F-критерия Фишера, применяемого в дисперсионном анализе, ими был использованы критерии критерия χ^2 и критерий χ^2 с поправкой Йетса? Или это как раз и говорит о том, что авторы использовали, например, двухвыборочный критерий Вилкоксона? При этом авторы дают ссылку на источник [10]: Bernhard V.M. Carotid artery stenosis. Association with surgery for coronary artery disease / Bernhard V.M., Johnson, W.D., Peterson J.J. // Arch. Surg. 1972. — Vol.105. — № 12. — P. 837-840. Однако у всех ли читателей эта статья, опубликованная 40 лет назад в 1972 г., лежит на рабочем столе? Если обратиться, например, на сайт самого журнала, где опубликована данная статья, то для ознакомления с её полным текстом редакция предлагает купить такой доступ. Чем же объяснить нежелание авторов статьи «Хирургическое лечение мультифокального атеросклероза» более полно изложить читателям методику анализа собранных ими данных? А ведь ясное и понятное изложение методики анализа этих данных может служить примером для подражания многим читателям этой статьи, в том числе и их коллегам, только начинающим свои научные изыскания.

Прочитай и перешли своим коллегам…

Более того, непонятно, как можно использовать дисперсионный анализ, если при наличии двух групп сравнения и некоторого количества различных осложнений очевидным методом сравнения частот является анализ таблиц сопряжённости? О том, что речь должна идти именно об анализе таблиц сопряжённости говорит и тот факт, что поправка Йетса используется как раз в анализе таблиц сопряжённости. Не сообщают авторы и о том, какую величину они размещают после знака «±» в выражениях вида 51,4±6,2. Внимательное прочтение статьи не обнаруживает также и конкретных результатов использования упомянутых выше методов – дисперсионного анализа, – ни критерия χ^2 ни критерия χ^2 с поправкой Йетса, и т.д. Т.е. нет ни самих значений статистических критериев, ни величин достигнутого уровня значимости для этих критериев. И вновь возникает вполне закономерный вопрос: с какой целью авторы статьи перечислили эти методы и статистические критерии, не приведя при этом никаких конкретных результатов их использования?

60) ВЕСТНИК НОВЫХ МЕДИЦИНСКИХ ТЕХНОЛОГИЙ – 2007 – Т. XIV, № 2 – С. 171. Маркеры активации тромбоцитов и их влияние на функциональное состояние шунтов после проведении операции аорто-коронарного шунтирования пациентам с эссенциальной артериальной гипертензией.

Таблица 1

Показатели 4ПФ у пациентов ИБС и АГ различной степени после операции КШ (M±m)

Группы	4ПФ (МЕ\мл)
Контроль	7,1±0,13
II с НХ	7,8±0,04 *1
III с НХ	8,5±0,07 *1,3
II с ГХ	9,1±0,05 *1,2
III с ГХ	10,2±0,08 *1,2,3

Здесь и далее: *1 – p<0,05 – по сравнению с контролем; *2 – p<0,05 – по сравнению с аналогичной группой; *3 – p<0,05 – по сравнению с АГ II ст.

В табл. 1 авторы приводят результаты сравнения для 5 групп. Авторы ничего не сообщают о том, с помощью каких конкретно статистических методов и критериев проводили сравнение между собой 5 групп. Поскольку в статье неоднократно приводятся выражения $p < 0,05$, то из этого можно сделать вывод о том, что авторы приняли в качестве критического уровня значимости величину p=0,05. Однако, учитывая тот факт, что авторы проводят сравнение между собой 5 групп, то в этом случае необходимо учесть проблему множественных сравнений. И в этом случае величина критического уровня значимости должна быть значительно понижена. В статье используются выражения, описывающие результаты сравнения групп и содержащие слово «достоверно», что, в данном контексте, некорректно [24]. Существенным недостатком данной публикации является также полное отсутствие в разделе «Материал и методы» информации об использованных авторами методах статистического анализа имеющихся данных. Как и во многих других статьях, авторы используют выражения вида 51,5 ± 0,5, не сообщая, однако, при этом, что за величины расположены в них после знака « ± ».

61) СИБИРСКИЙ МЕДИЦИНСКИЙ ЖУРНАЛ № 3, 2009 (выпуск 2), с. 81-87. Диагностическое значение неврологических и нейропсихологических симптомов и признаков в зависимости от длительности искусственного кровообращения и среднего артериального давления.

В начале раздела «МАТЕРИАЛ И МЕТОДЫ» авторы сообщают: «Для оценки особенностей реакции мозга и формирующихся постишемических изменений мы использовали в качестве клинической модели результаты обследований пациентов, перенесших экстракорпоральное кровообращение (ИК) продолжительностью свыше 120 мин (M ± m = 169,4±55,30) при поверхностной (34-33°С) гипотермии тела».

Прочитай и перешли своим коллегам…

Статистические методы обработки полученных данных включали следующие: обработка результатов методами вариационной статистики – определяли значения среднего арифметического (M), стандартного отклонения (a), а также ошибку среднего арифметического (т). Для статистической обработки полученных количественных данных в двух группах мы пользовались критерием t Стьюдента. При статистической обработке качественных показателей в двух группах использовался критерий χ^2. При анализе значимости показателя взаимной сопряженности (он же является коэффициентом корреляции Пирсона т) использован в адаптированной методике тест на линейную ассоциацию Mantel-Haensxel. Статистически значимыми считались различия между группами, когда расчетное значение t и χ^2 превышало их критическое значение по таблице при $p<0,05$. Вычислительные алгоритмы факторного анализа и анализа взаимной сопряженности качественных признаков были выполнены с использованием метода дисперсионного анализа (однофакторная ANOVA-МОДЕЛЬ) и анализа качественных признаков (коэффициенты линейной ассоциации и взаимной сопряженности). На каждого пациента составлялись таблицы, куда заносились балльные (количественные) параметры неврологического и нейропсихологического статуса. На основании полученных данных вычислялись статистическими и вероятностно-математическими методами: а) статистическая значимость различий; б) корреляции; в) дисперсионный анализ; г) значимость (или вес) симптома или признака в патологическом процессе, его квалификация.

Слева приведён фрагмент раздела «МАТЕРИАЛ И МЕТОДЫ», в котором авторы описывают используемые в работе статистические методы анализа. В нём авторы сообщают, что они «… определяли значения среднего арифметического (M), стандартного отклонения (a), а также ошибку среднего арифметического (т)».

Однако нигде не сказано, что за величина обозначается буквой «m» в выражениях M ± m. Что это – ошибка авторов, или же ошибка технических работников редакции ПКК, которые, возможно, не имея электронной авторской версии статьи, набирали её вручную? Авторы также сообщают, что использовали критерий Стьюдента для сравнения количественных данных в двух группах. В двух таблицах авторы приводят некие результаты статистического анализа, не упоминая при этом, какими конкретно статистическими методами они были получены. Например, вот как выглядит табл. 1 (см. справа).

Таблица 1

Зависимости между неврологическими, нейропсихологическими симптомами и факторами искусственного кровообращения (n=30)

№ п/п	Неврологический симптом, признак	Неврологический статус, НП статус, ИК-фактор	Значимость P(к)
1.	НВ	САД 70 4	p<0,01
2.	нв + нп	САД 70	p<0,01
3.	Нарушения сна	НП	<0,01
4.	Мозжечковая атаксия	НП	<0,01
5.	Мозжечковая атаксия	НВ	<0,05
6.	Центральный парез VII D	1 мин.	<0,01
7.	Брадикинезия	САД 70 1	<0,05

Видим, что в последнем столбце авторы приводят выражения вида p<0,01, <0,05. Данная таблица предваряется следующей информацией: «Для получения ответов на ряд вопросов, связанных с оценкой информативности, частоты встречаемости и достоверности симптомов, признаков или тестов, их значения (или веса) в используемом при ОГИМ клиническом анализе, мы провели факторный анализ (дисперсионный анализ, однофакторная ANOVA-модель)». Обратим внимание, что использование дисперсионного анализа, который авторы называют факторным анализом (см. http://www.biometrica.tomsk.ru/factor.htm), имеет свои ограничения. Эти ограничения идентичны ограничениям при использовании параметрического t-критерия Стьюдента. В частности, это наличие нормального распределения количественного признака ВО ВСЕХ ГРУППАХ СРАВНЕНИЯ, а также равенство генеральных дисперсий этого количественного признака также ВО ВСЕХ ГРУППАХ СРАВНЕНИЯ. Видимо авторы данной статьи не подозревают о существовании данных ограничений, и по этой причине не упоминают о проверке этих ограничений. Между тем, количественные признаки, оцениваемые в медицине, редко имеют нормальное распределение. Этот факт достаточно подробно описан нами в разделе «СТАТИСТИЧЕСКАЯ ВАМПУКИЗАЦИЯ, ОНА ЖЕ ВСЕОБЩАЯ СТЬЮДЕНТИЗАЦИЯ» [42]. (http://www.biometrica.tomsk.ru/kuzbass5.htm) Столь же редко встречается и равенство дисперсий. В результате одновременное выполнение обеих этих ограничений встречается в нескольких процентах случаев. Поэтому отсутствие проверки авторами статьи этих условий не вызывает доверия к выводам, приведённым в табл. 1-2.

Другой аспект результатов данного исследования заключается в том, что авторы применили методы сравнения средних для 2 и более групп к величинам, которые в принципе не могут иметь нормального распределения. Речь идёт о балльных признаках. Вот как пишут об этом авторы данной статьи: «На каждого пациента составлялись таблицы, куда заносились балльные (количественные) параметры неврологического и нейропсихологического статуса. ….

Прочитай и перешли своим коллегам…

Использовалась 4-балльная шкала оценок нарушений ВПФ по А.Р. Лурия». Поэтому для балльных признаков должна использоваться не процедура классического однофакторного дисперсионного анализа – ANOVA, а процедуры непараметрического сравнения групп. Отметим, что в непараметрическом анализе используется несколько статистических критериев, например, критерий Ван дер Вардена, медианный критерий, и т.д.

Далее авторы пишут: «При анализе значимости показателя взаимной сопряжённости (он же является коэффициентом корреляции Пирсона r) использован в адаптированной методике тест на линейную ассоциацию Mantel-Haensxel». При этом авторы не дают никакой ссылки на литературный источник с описанием этой адаптированной методики. Между тем, линейная корреляция Пирсона используется для изучения взаимосвязи двух непрерывных, количественных признаков. Тогда как взаимная сопряжённость по Мантелю–Ханзелю относится к анализу таблиц сопряжённости, в которых анализируется взаимосвязь двух качественных, либо ранговых признаков. И по этой причине неясно, о какой методике, связывающей воедино два совершенно различных вида статистического анализа, идёт речь в статье. Отметим также неверное написание фамилии одного из авторов упоминаемого теста. Вот как записываются их фамилии в одной из статей: Mantel, N. and Haenszel, W. (1959), "Statistical Aspects of the Analysis of Data from Retrospective Studies of Disease," Journal of the National Cancer Institute, 22, 719–748.

Диагноз: патологическая статистика в кардиологии

Никакую проблему нельзя
решить на том же уровне,
на котором она возникла.

Альберт Эйнштейн

Сравнивая содержание первого и второго разделов данного обзора, можно констатировать, что в статьях ПКК очень часто встречается как некачественное описание статистики, так и непосредственно некачественное её использование. Однако факт наличия в статьях некачественного описания статистики не обязательно означает, что некачественно был проведён и сам этап статистического анализа кардиологических данных. Возможно, что статистический анализ с использованием несложных статистических процедур был выполнен авторами правильно. Например, при сравнении средних двух групп с помощью t-критерия Стьюдента проверялись как гипотеза нормальности в обеих группах, так и гипотеза равенства генеральных дисперсий. Однако при написании раздела «Материал и методы» авторы не сочли необходимым описать эти детали. Мотивы такой асимметрии могут быть разными. Однако более вероятно, что некачественное описание статистического анализа является проявлением такого же уровня и статистического анализа кардиологических данных.

Ниже перечислим наиболее часто встречающиеся некорректности описания и выполнения статистического анализа в проанализированных выше статьях.

1. Использование выражений типа «M ± m» и «M ± δ», без указания, что за величины соединены знаком « ± ».
2. Использование выражений типа M ± m для балльных признаков.
3. Использование t-критерия Стьюдента без проверки двух ограничительных условий.
4. Использование t-критерия Стьюдента для сравнения средних более чем в двух группах, без проверки двух ограничительных условий.

Прочитай и перешли своим коллегам…

5. Использование классического дисперсионного анализа (ANOVA) без проверки двух ограничительных условий.
6. Использование t-критерия Стьюдента к дискретным, балльным признакам.
7. Использование в тексте статей конкретных значений достигнутого уровня статистической значимости, без указания статистических критериев, для которых они вычислялись.
8. Упоминание о том, что уровень значимости принимался более 5%.
9. Упоминание о том, что уровень значимости принимался более 95%.
10. Использование выражений о результатах проверки статистических гипотез со словом «достоверно».
11. Упоминание о применении многомерных методов статистического анализа без конкретизации этих методов и приведения результатов их использования.
12. Использование t-критерия Стьюдента в среде пакета EXCEL, в котором нет процедуры проверки нормальности распределения.
13. Использование корреляционного анализа без конкретизации используемого коэффициента корреляции.
14. Использование выражений «непараметрические данные», «непараметрические признаки», «параметрические признаки».
15. Упоминание в разделе «Материал и методы» конкретных методов статистического анализа, результаты использования которых, вообще не приводятся в тексте статей.
16. Использование выражения вида «$p > 1$», в которых, судя по контексту, величина «p» является вероятностью.

Общеизвестно, что для эффективного лечения любого заболевания необходима обязательная реализация следующих условий:

1) Постановка правильного диагноза;

2) Установление причин возникновения заболевания;

3) Формулировка конкретных методов лечения конкретного больного (выбор лекарственных препаратов, лечебных процедур, и т.п.).

Перечисленные выше признаки "заболевания" позволяют поставить диагноз "статистическая патология". Полагаю, что смысл этого термина понятен и очевиден.

Констатируемая ситуация может вызываться причинами разного уровня. В данном разделе обсудим наиболее простые причины первого уровня. Более же глубокие основания появления в т. ч. и этих причин, обсудим в следующем разделе обзора.

Итак, во-первых наблюдаемая статистическая патология есть результат выполнения этапа статистического анализа самими медиками, не имеющими соответствующей профессиональной подготовки в области биостатистики. Недостаточный уровень владения методами биостатистики у медиков вполне объясним. Ведь медицинские вузы готовят врачей, а не специалистов в области биостатистики. Т.е. основной целью обучения в медицинских вузах является подготовка специалиста, способного лечить больных, а не заниматься научными исследованиями, в т. ч. с помощью сложных статистических методов. Вполне естественно, что даже в научных медицинских организациях на первом плане должна стоять задача лечения пациентов. Однако отсутствие в них служб статистического сервиса есть признак того, что научные исследования в перечне доминирующих целей находятся в самом конце. В результате исследователи в этих

организациях вынуждены проводить статистический анализ самостоятельно на непрофессиональном уровне, допуская при этом массу ошибок и некорректностей. В итоге возникает весьма удручающая асимметрия: отличные специалисты в конкретной области медицины, проводя сбор ценнейших данных по результатам своей практической и исследовательской деятельности, обесценивают эти исходные данные, теряя при этом информацию, которая по уровню ценности и будущей востребованности равна уровню их квалификации в данной области медицины.

Более того, получаемая в результате некорректного статистического анализа информация является подчас ошибочной, и тем самым способной наносить вред пациентам, к которым будет она применяться. Вот как об этом написано в статье [18]. «Когда в статье нет указания на критерий для расчета P, вернее, искомых различий, когда указывая критерий не говорят о параметрах необходимых для его применения (распределение в популяции или в выборке) тогда P и становится пресловутым мерилом метафизики, а не науки». Итак, отсутствие в организациях занятых как лечением больных, так и наукой, служб специального статистического сервиса (лабораторий биостатистики) означает минимизацию важности научной деятельности со стороны руководства этих организаций.

У врача основным объектом его деятельности является пациент. Тогда как у журнала таким «пациентом» является читатель. Врачебная практика во всём мире не обходится без врачебных ошибок. Их последствия имеют в разных странах разные объёмы. Так согласно статистике Всемирной организации здравоохранения каждый год в мире от врачебных ошибок гибнет людей больше, чем в результате ДТП. «По данным бывшего Минздравсоцразвития, готовившего проект закона о страховании гражданской ответственности медицинских организаций перед пациентами, смертность от врачебных ошибок в России наступает в 45 – 50 тыс. случаях ежегодно, – говорит директор Центра медицинского права юрист Алексей Панов» (http://newvesti.info/ezhegodno-ot-vrachebnyx-oshibok-v-rossii-umirayut-okolo-100-tysyach-chelovek/) Законодательство, регламентирующее ответственность за врачебные ошибки, было ещё в глубокой древности. Наиболее известным из таких законодательных актов, является высеченный на камне Свод законов (законник) Хаммурапи, который правил Вавилоном и объединенной Месопотамией в 1792-1750 гг. до н. э. Однако в любом случае для доказательства факта врачебной ошибки необходимо установление с помощью экспертизы причинно-следственной связи. Очень часто всего такие связи достаточно очевидны. Поскольку последствия таких ошибок имеют небольшую протяжённость, как во времени, так и в пространстве.

Совершенно иное положение во взаимодействии авторских ошибок в научных публикациях, и читателей. Здесь гораздо большее количество читателей, и больше рассеяние контактов «публикация – читатель» как во времени, так и в пространстве. Более того, очевидно, что любая авторская ошибка имеет самую минимальную вероятность нанесения каких-то нежелательных последствий конкретным пациентам через читателя. Таким образом, последствия авторских ошибок в медицинских публикациях сказываются в основном на экономической эффективности проводимых научных исследований.

В 2007 г. в «Международном журнале медицинской практики» (вып. 2, с. 19-35) была опубликована наша статья «Ошибки статистического анализа биомедицинских данных» (http://www.biometrica.tomsk.ru/error.htm). Процитируем фрагмент этой статьи.

Прочитай и перешли своим коллегам…

«Чем же можно объяснить такую терпимость к многолетнему воспроизводству статистических ошибок в российских публикациях биомедицинской тематики? Основным барьером на пути уменьшения многочисленных ошибок в использовании статистических методов, является некритическое отношение таких специалистов, как к собственной исследовательской практике, так и к практике своих коллег. Результатом этого является неявное следование следующим установкам:

- использовать тот статистический метод, который предложил руководитель (ослушаться нельзя, а если кто-то обвинит в некорректности, то виноват шеф);
- применять данный метод потому, что он есть в доступном статистическом пакете (надо же написать в статье, что использовал статистический пакет);
- использовать те методы, которые ранее уже были самим автором использованы (за них никто не поругал, значит можно и впредь их применять).
- использовать данный метод потому, что его уже использовал другой автор опубликованной статьи (диссертации) по этой же тематике (раз статью опубликовали (диссертацию защитили), значит можно использовать, ведь в редакциях (диссертационных советах) не дураки же сидят).

Нередко следование этим установкам приводит к генерации цепочек так называемых «мемов», когда несколько авторов последовательно заимствуют друг у друга одну и ту же ошибку. (Меметический анализ описаний методов статистики http://www.biometrica.tomsk.ru/lis/index19.htm)

Можно выделить следующие причины появления и живучести таких ошибок:
Отсутствие знаний по статистике у авторов публикаций, членов журнальных редколлегий и диссертационных советов, что не позволяет им осознать всю степень ошибочности выводов в таких публикациях.

1. Деформация цели публикации. Нередко истинной целью публикации является не стремление ученого обнародовать, легализовать надежные, достоверные результаты исследования, а лишь сам факт быстрейшей публикации даже ненадежных выводов, как доказательство «научности» результатов.
2. Отдаленность во времени и пространстве возможных негативных последствий опубликования сомнительных результатов от автора, времени и места публикации.
3. Ограниченность доступа широкой читательской аудитории к медицинским публикациям, содержащим ошибочные выводы. Это связано как с ограниченными тиражами медицинской периодики, так и с ограничением доступа к медицинским диссертациям. В ряде медицинских вузов доступ к защищенным диссертациям разрешается ректором вуза.
4. Отсутствие доступа для читателей к исходным данным автора публикации. Автор, зная, что любой читатель, сомневающийся в истинности авторских выводов, имеет возможность немедля (или же спустя некоторое время) получить эти данные, и самостоятельно перепроверить авторские выводы, будет совершенно иначе относиться к процедуре статистического анализа этих данных.
5. Незаинтересованность редакций периодических изданий, ректоров медицинских вузов, диссертационных советов, ВАК РФ и дирекций НИИ в повышении качества журнальных публикаций и диссертаций».

Одной из важных причин такого состояния в медицинских исследовательских центрах является отсутствие общероссийских отраслевых стандартов, определяющих требования к публикациям в биомедицинских журналах. Т.е. нет таких стандартов ни в РАН, ни в РАМН, ни в Минздраве, ни в Минобрнауке, ни в ВАК, и т.д. Рассмотрим, например, «Правила оформления

статей для авторов журнала «ПАТОЛОГИЯ КРОВООБРАЩЕНИЯ И КАРДИОХИРУРГИЯ» (ПКК, № 3 2010).

«Авторы предоставляют статьи на русском языке, **распечатанными на бумаге форматом А4 и в электронной версии. Размер кегля 14 для Times New Roman, межстрочный интервал 1,5** (*выделено нами. В.Л.)***.** Объем материалов: оригинальной статьи 10–12 страниц, обзоры, лекции 12–15 страниц, клиническое наблюдение 4–5 страниц. Кроме того, на отдельных страницах – список литературы (в оригинальных статьях цитируется не более 15 источников, в обзорных – не более 45), таблицы (не более трех), рисунки (не более трех) и подписи к ним. Рукопись должна быть выверена и датирована. На титульной странице необходимо указать:
1) индекс универсальной десятичной классификации (УДК),
2) заглавие публикуемого материала,
3) инициалы и фамилии авторов. Авторский коллектив может состоять не более чем из восьми человек. Если авторы работают в разных учреждениях, необходимо знаками * и ** отметить это,
4) полное наименование учреждения (учреждений), в котором (которых) была выполнена работа,
5) почтовый адрес учреждения,
6) контактные телефоны,
7) адрес электронной почты (сообщается читателям),
8) аннотацию на русском языке, которая в сжатой форме отражает существо излагаемого вопроса, материал и методы исследования, результаты работы и выводы, содержит не более 200–250 слов,
9) ключевые слова (не более 3–5) на русском языке; отделяются друг от друга точкой с запятой.
На английском языке предоставляются:
1) заглавие публикуемого материала,
2) инициалы и фамилии авторов,
3) полное наименование учреждения (учреждений), в котором (которых) была выполнена работа,
4) почтовый адрес учреждения,
5) аннотация,
6) ключевые слова (не более 3–5); отделяются друг от друга точкой с запятой. В конце статьи необходимо указать сведения об авторах: фамилию, имя, отчество (полностью), ученую степень и звание (полностью), должность и место работы (полностью). Рукопись должна быть подписана всеми авторами, берущими на себя тем самым **ответственность за достоверность** предоставленных материалов. После поступления в редакцию все статьи отсылаются на внешнее рецензирование двум рецензентам. При получении положительных рецензий **решение о принятии статьи к публикации выносится на основании** ее значимости, оригинальности, **достоверности предоставленных материалов после экспертного совета членов редакционной коллегии журнала**. Редакция предоставляет авторам рецензии рукописей по требованию, оставляет за собой право отклонять статьи в случае получения на них отрицательных рецензий, а также вносить редакторскую правку. В случае отказа в публикации редакция направляет авторам мотивированный отказ.

Требования к оформлению основного текста и иллюстративных материалов
Статья: основной текст, список литературы, иллюстрации, подписи к рисункам и таблицам – предоставляется в редакцию **распечатанной на бумаге форматом А4 и в электронной версии, выполненной в текстовом редакторе Microsoft Word. Размер кегля 14 для Times New Roman, межстрочный интервал 1,5.** (*выделено нами, поскольку данное требование уже было в самом начале. В.Л.*) Страницы нумеруются (внизу справа). Не следует производить табуляцию, разделять абзацы пустой строкой, использовать макросы, сохранять текст в виде шаблона и с установкой «только для чтения», форматировать текст и расставлять принудительные переносы. Разделы

оригинальной статьи: введение, **материал и методы**, результаты, обсуждение, выводы (заключение), список литературы. Разделы Введение и Заключение выделять заголовками не следует. В разделе **Выводы** выводы представлять нумерованным списком. Допускается не более трех таблиц и трех рисунков в одной статье. Таблицы и иллюстрации вставлять в текст не нужно, они помещаются в отдельных файлах (озаглавленных соответственно tabl., ris.). В тексте обязательно должны быть ссылки на все таблицы. Место, где в тексте должна быть помещена таблица, отмечается пропуском строк и указанием номера таблицы. Каждая таблица должна иметь заголовок. В тексте обязательно должны быть ссылки на все иллюстрации. Иллюстративный материал (черно-белые и цветные фотографии, рисунки, диаграммы и графики) обозначать как «рис.» и нумеровать в порядке их упоминания в тексте. Место, где в тексте должен быть помещен рисунок, отмечается пропуском строк и указанием номера рисунка. На обороте рисунка карандашом указать его номер и верх рисунка. Допустимо представление графических черно-белых рисунков, выполненных на компьютере в программе Corel Draw; все надписи и символы на них следует перевести в кривые. Электронный вариант рисунков прилагать обязательно (разрешение не менее 300 dpi, формат .tif). Просим вас не изменять исходный электронный формат создаваемого вами графического объекта (фотографии, эхограммы и пр.). Диаграммы и графики должны быть выполнены в формате MS Excel и распечатаны на лазерном принтере. К ним необходимо приложить исходные данные, по которым они были построены, и электронный вариант. Рисунки, графики и диаграммы должны быть только черно-белыми, без цветных элементов и мелких (сплошных) заливок.

Библиографические ссылки

Библиографические ссылки в тексте статьи даются номерами в квадратных скобках в соответствии с пристатейным списком литературы. В оригинальных статьях цитируется не более 15 источников, в обзорных – не более 45. В список литературы не включаются неопубликованные работы и учебные пособия. Пристатейный список литературы должен оформляться следующим образом. Источники указываются строго в алфавитном порядке; сначала работы авторов на русском языке, затем на других языках. Все работы одного автора нужно указывать по возрастанию годов издания. В библиографической ссылке на журнальную статью следует указать ее авторов, через // дать название журнала, год выпуска, том, номер, обязательно страницы, на которых статья опубликована. **Название статьи указывать не нужно**. Авторы несут ответственность за правильность данных, приведенных в пристатейном списке литературы. Рукописи, оформленные не в соответствии с указанными правилами, не рассматриваются. Корректура авторам не высылается, и вся дальнейшая сверка проводится редакцией по авторскому оригиналу. Направление в редакцию работ, которые уже опубликованы в иных изданиях или посланы для публикации в другие редакции, не допускается. Не принятые к печати рукописи авторам не возвращаются. Срок хранения рукописи – 2 года. Статьи в журнале «Патология кровообращения и кардиохирургия» публикуются бесплатно, в том числе статьи соискателей-аспирантов».

Приведённые выше редакционные Требования состоят содержат 770 слов. В них, в частности, упоминаются и необходимые разделы рукописи статьи, в т.ч. и раздел «Материал и методы». Однако в самих требованиях нет ни единого слова относительно описания статистических методов, которые использовали авторы исследования. В частности, нет ни одного слова о том, что авторы обязаны описывать использованные ими статистические методы, объяснять выражения типа M ± m, задавать значения критического уровня статистической значимости, приводить величины достигнутых уровней значимости, аргументировать

корректность использования тех или иных статистических методов и критериев, и многие другие детали статистического анализа. Этим самым редколлегия ПКК разрешает авторам рукописей либо вообще никак не описывать статистические аспекты исследования, либо делать это так, как они считают и умеют это делать. Например, ограничиваться лишь упоминанием об использовании статистических методов, но при этом никак не конкретизировать детали этого этапа исследования. Либо вообще игнорировать описание статистических аспектов исследования. Отметим, что при этом требования к описанию иллюстраций содержат 99 слов. И в этих требованиях уже несколько лет используется ошибочное написание названия пакета EXCEL, которое записано как Excell, как в последней версии редакционных требований, так и в 3-х предыдущих версиях, опубликованных в 4-м номере ПКК за 2010 г., 4-м номере ПКК за 2011 г., и 4-м номере ПКК за 2012 г. Возникает вопрос, неужели за все эти годы никто из членов редколлегии и авторов не заметил этой ошибки? Или же никто из них вообще не читал эти требования, а составлялись они кем-то, кто весьма далёк от работы с программой EXCEL? Или же здесь нет никакой ошибки, и речь идёт некоем малоизвестном продукте Excell? Сделав запрос в поисковике Яндекс, найдём следующую информацию об Excell.

1. Катушка DETECH Excell 5" для Minelab Exp/E-Trac. Разработана для точечного поиска на сильно замусоренных участках. При использовании этой поисковой катушки вы поймете, как много ещё находок можно сделать в местах еще недавно считавшихся отработанными. Стоимость 1 катушки составляет 4900 рублей.

2. Аккумулятор EXIDE Excell EB620 12V 62Ah 540A R+. Аккумуляторная батарея, предназначенная для уверенного пуска двигателей и обеспечения энергией электрического оборудования автомобилей со средним уровнем оснащённости. Стоимость 1 экз. 3070 рублей.

И т.д. Однако ни одна из найденных ссылок не содержит ничего, что позволяло бы предположить, что именно об этом уникальном предмете, и написано в редакционных требованиях. Более того, непонятен смысл следующей фразы из этих требований: "Иллюстрации следует присылать в исходной программе, например, Excell для диаграмм и графиков". Т.е. эти иллюстрации следует присылать С САМОЙ ИСХОДНОЙ ПРОГРАММОЙ Excell? В действительности видимо подразумевается здесь нечто иное. Например, можно предположить, что иллюстрации следует присылать в виде ФАЙЛОВ ФОРМАТА ПАКЕТА EXCEL. Однако и в этом случае возникает вопрос, о каком конкретно формате файла идёт речь? Ведь пакет EXCEL имеет разные версии, и в них также есть и разные форматы файлов, содержащих иллюстрации.

Для сравнения, на сайте известного журнала JAMA редакционные требования (версия от 26.11.2013) (http://jama.jamanetwork.com/public/instructionsForAuthors.aspx) содержат более 13 тысяч слов. В том числе только один раздел «Manuscript Checklist», включающий 21 пункт, содержит 650 слов. В этих требованиях содержатся в т. ч. и такие важные моменты, как **обязательное указание автора, который берёт на себя ответственность за полноту данных и точность анализа этих данных. Там же говорится и о том, что анализ данных должен проводиться независимым биостатистиком, имеющим**

соответствующий уровень знаний и опыт анализа данных, и т.д. Ещё больший объём занимают редакционные требования журнала BMJ, издающегося, как известно, уже более 170 лет.

Если рассмотреть взаимосвязь редакционных требований, и описания статистических аспектов в статьях ПКК, то невозможно сказать, что является причиной, а что – следствием. Ведь отсутствие в редакционных требованиях статистических аспектов публикаций является отражением представления членов редакционной коллегии и редакционного совета (http://www.meshalkin.ru/files/magazine/mag_redaction.htm) о структуре своих собственных публикаций в ПКК. Поскольку анализ авторского состава статей в ПКК, показывает, что **наибольшая частота публикаций в ПКК наблюдается как раз у членов редколлегии и редакционного совета**. Которые, в силу медицинского образования и большой занятости как управленческой деятельностью, так и лечебной практикой, не могут обладать необходимыми знаниями по биостатистике. Поэтому вероятнее всего оба этих проявления являются следствиями одних и тех же причин.

Отсутствие требований к описанию статистики, подобных тем, что приведены в книге «Как описывать статистику в медицине: руководство для авторов, редакторов и рецензентов» (Т.А. Ланг, М. Сесик; пер. В. П. Леонов. – М.: Практическая медицина, 2011. – 480 с.), создаёт проблемы не только для отечественных публикаций. Летом 2012 г. на заседании Учёного совета Томского НИИ кардиологии ТНЦ СО РАМН учёные подняли вопрос о том, что зарубежные журналы не публикуют их статьи, поскольку в них некорректно использованы методы статистического анализа. Автор этих строк, будучи несколько лет членом редколлегии издаваемого в этом НИИ «Сибирского медицинского журнала», неоднократно обращался к директору НИИ с предложением о создании в НИИ кардиологии лаборатории биостатистики. Однажды директор НИИ согласился с этим предложением, и даже рассказал о том, где будет находиться эта лаборатория в новом корпусе, который в то время ещё строился. Однако после окончания строительства этого корпуса вместо лаборатории биостатистики на этой площади была открыта … церковь.

В качестве примера-антипода, содержащего достаточно подробные описания статистики, приведём **копию раздела «Методы статистического анализа»** статьи «Сравнительная эффективность применения высокой и низкой дозы лозартана у больных с сердечной недостаточностью: результаты исследования HEAAL» (Heart failure Endpoint evaluation of Angiotensin II Antagonist Losartan) Источник: Konstam M.A., Neaton J.D., Dickstein K., et al. Effects of high-dose versus low-dose losartan on clinical outcomes in patients with heart failure (HEAAL study): a randomised, double-blind trial. **Lancet 2009**; 374 (9704): pp. 1840—1848) из выпуска №1 за 2010 г. (с. 4–11) опубликованную в журнале «Доказательная кардиология», основанного в 2003 г. (http://www.mediasphera.ru/journals/dokcard/about/). Напомним, что одним из учредителей этого журнала является Межрегиональное общество специалистов доказательной медицины. В этом журнале публикуются материалы из международных медицинских изданий с комментариями: обзоры наиболее важных событий в кардиологии, критический анализ результатов наиболее значимых рандомизированных контролируемых испытаний, результаты вторичного анализа данных, полученных в ходе медицинских исследований, а также результаты систематических

обзоров, мета-анализов и публикаций, посвященных оценке экономических аспектов различных вмешательств. Специальный раздел журнала посвящен анализу новых клинических рекомендаций. Итак, вот раздел «Методы статистического анализа» данной статьи, содержащий 423 слова.

«Методы статистического анализа
Анализ основного и дополнительных показателей выполняли исходя из допущения, что все больные получали назначенное лечение. Для сравнения частоты развития тяжёлых осложнений, включённых в основной и дополнительные показатели, при использовании лозартана по 150 и 50 мг/сут применяли методы анализа, основанные на оценке времени до развития определенного клинического исхода (анализ кривых Каплана—Мейера и модели пропорциональных рисков Кокса). Данные о больных, с которыми был потерян контакт или которые не наблюдались до 31 марта 2009 г., были включены в окончательный анализ как незавершённые наблюдения. Сравнение эффективности применения лозартана по 150 и по 50 мг/сут оценивали с помощью моделей Кокса с учётом таких факторов, как применение β-блокаторов, географический регион и характер исследуемой терапии. Выделяли следующие географические регионы: Европа, Средний Восток и Африка; Азия и Тихоокеанский регион; Латинская Америка. Проверку допущений для отношений рисков выполняли путём анализа взаимодействия между учитываемыми факторами и продолжительностью наблюдения, представленного в логарифмической форме. Для сравнения общего числа госпитализаций по поводу утяжеления СН между группами применяли модели Кокса. Основной показатель анализировали также в подгруппах больных с заранее определёнными демографическими характеристиками и исходными данными. Гетерогенность отношений рисков оценивали для каждой подгруппы путем включения в расширенную модель Кокса данных о взаимодействии между применяемой терапией и определённой подгруппой. Изменения АД, скорости клубочковой фильтрации, которую рассчитывали с помощью формулы MDRD (Modification of Diet in Renal Disease), и концентрации калия в крови по сравнению с исходными данными в каждой из групп оценивали с помощью ковариационного анализа с использованием таких исходных ковариат, как изучаемый показатель, регион проживания и применение β-блокаторов. Долю больных, у которых через 12 мес отмечалось увеличение концентрации калия в крови более 6,0 ммоль/л и креатинина не менее чем в 2 раза по сравнению с исходной, оценивали с помощью критерия χ2 Мантеля—Хенцеля с использованием стратификации в зависимости от региона проживания и применения β-блокаторов. Сравнение изменений ФК по классификации NYHA между группами проводили с помощью критерия Вилкоксона. Частоту развития побочных эффектов и полного прекращения приёма исследуемых препаратов, как в целом, так вследствие определённых причин, сравнивали с помощью методов, основанных на определении времени до развития таких исходов. Прогностическую ценность исходных показателей как факторов риска развития побочных эффектов оценивали с помощью моделей Кокса. В ходе такого анализа не учитывали случаи прекращения приема препаратов в течение 14 дней до смерти, за исключением тех случаев, когда причиной прекращения терапии были побочные эффекты. При выполнении анализа частоты развития побочных эффектов данные о больных, которые прекратили прием исследуемых препаратов, были включены в окончательный анализ как незавершённые наблюдения; при этом использовали данные, полученные через 14 дней после прекращения приема препарата. Все виды статистического анализа проводили с помощью пакета программ SAS версия 9.1».

Зададимся вопросом, является ли отсутствие в журнале ПКК требований об описании статистических аспектов исследования столь же уникальным явлением, каким уникальным является и сам НИИ им. акад. Е.Н. Мешалкина? Для ответа на этот вопрос приведём краткие описания редакционных требований ряда отечественных кардиологических журналов, список которых приведён на сайте «Российского кардиологического журнала»: «Журналы, публикующие статьи по специальности – кардиология». (http://roscardio.ru/ru/recommendations-for-the-articles-rjc/item/166-gurnali-publikuuwie-stati.html)

Прочитай и перешли своим коллегам…

Вот эти журналы:

1. "Электронный журнал ANGIOLOGIA.ru" (http://www.angiologia.ru/journal_angiologia/publish/) Редакционные требования содержат 597 слов. Нет ни одного слова об описании статистики.

2. "Кардиология и сердечно-сосудистая хирургия / Cardiology & cardiovascular surgery" (http://www.mediasphera.ru/journals/cardsurg/pravila/) Редакционные требования содержат 825 слов. Нет ни одного слова об описании статистики. В требованиях читаем: «Все математические формулы должны быть тщательно выверены и **чётко вписаны черными чернилами**». То есть для членов редколлегии важнее указать авторам рукописей цвет чернил, нежели привести требования по описанию статистики.

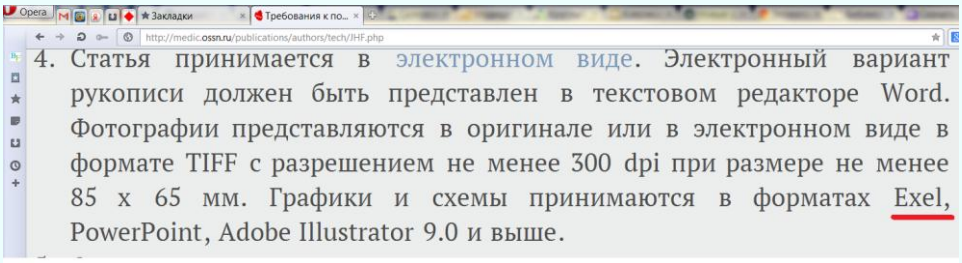
3. Журнал "Ангиология и сосудистая хирургия" (http://www.angiolsurgery.org/magazine/requirements/) Редакционные требования содержат 627 слов. Нет ни одного слова об описании статистики.

4. Журнал "Атеросклероз и дислипидемии" (http://noatero.ru/noa/pravila-dlya-avtorov) Редакционные требования содержат 555 слов. Нет ни одного слова об описании статистики.

5. Журнал "Кардиология" (http://www.cardio-journal.ru/ru/page/rules.html) Редакционные требования содержат 763 слова. Нет ни одного слова об описании статистики. Напомним, что 15 лет назад именно в этом журнале и была опубликована наша статья «Применение методов статистики в кардиологии (по материалам журнала «Кардиология» за 1993–1995 гг.) [44]. В ней мы достаточно подробно описали крайне низкий уровень описания статистики в 426 статьях данного журнала. Однако редколлегия журнала "Кардиология" в течение 15 лет так и не сочла необходимым и полезным для повышения качества кардиологических исследований изменить свои требования к рукописям в части описания статистики.

6. Журнал "Кардиология и сердечно-сосудистая хирургия / Cardiology & cardiovascular surgery" (http://www.mediasphera.ru/journals/cardsurg/pravila/) Редакционные требования содержат 825 слов. Нет ни одного слова об описании статистики.

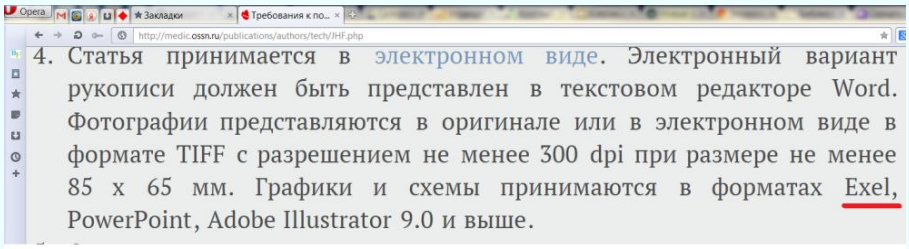
7. "СЕРДЦЕ: ЖУРНАЛ ДЛЯ ПРАКТИКУЮЩИХ ВРАЧЕЙ" (http://medic.ossn.ru/publications/authors/tech/Heart.php) Редакционные требования содержат 614 слов. Нет ни одного слова об описании статистики. В этих требованиях читаем следующее:


4. Статья принимается в электронном виде. Электронный вариант рукописи должен быть представлен в текстовом редакторе Word. Фотографии представляются в оригинале или в электронном виде в формате TIFF с разрешением не менее 300 dpi при размере не менее 85 х 65 мм. Графики и схемы принимаются в форматах Exel, PowerPoint, Adobe Illustrator 9.0 и выше.

Редколлегия журнала "СЕРДЦЕ: ЖУРНАЛ ДЛЯ ПРАКТИКУЮЩИХ ВРАЧЕЙ" состоит из 47 профессоров, д.м.н., среди которых немало академиков РАМН. Однако никто из

них не знает, что правильное название упомянутого выше пакета не **Exel,** а **Excel.** Поэтому можно сделать вывод о том, что в редколлегии «СЕРДЦЕ: ЖУРНАЛ ДЛЯ ПРАКТИКУЮЩИХ ВРАЧЕЙ» есть недостаточность знаний в области статистики.

8. Журнал "СЕРДЕЧНАЯ НЕДОСТАТОЧНОСТЬ" (http://medic.ossn.ru/publications/authors/tech/JHF.php) Редакционные требования содержат 642 слова. В том числе имеются следующие требования: «Указывается метод рандомизации пациентов и методы статистического анализа». Там же читаем следующее:


4. Статья принимается в электронном виде. Электронный вариант рукописи должен быть представлен в текстовом редакторе Word. Фотографии представляются в оригинале или в электронном виде в формате TIFF с разрешением не менее 300 dpi при размере не менее 85 х 65 мм. Графики и схемы принимаются в форматах Exel, PowerPoint, Adobe Illustrator 9.0 и выше.

Редколлегия журнала "СЕРДЕЧНАЯ НЕДОСТАТОЧНОСТЬ" состоит из 47 профессоров, д.м.н., среди которых немало академиков РАМН. Однако никто из них не знает, что правильное название упомянутого выше пакета не **Exel,** а **Excel**. Поэтому можно сделать вывод о том, что в редколлегии «Сердечной недостаточности» есть недостаточность знаний в области статистики.

9. Журнал "Профилактическая медицина / Preventive Medicine" (http://www.mediasphera.ru/journals/prof/pravila/) Редакционные требования содержат 707 слов. Нет ни одного слова об описании статистики.

10. "Кардиоваскулярная терапия и профилактика" (http://www.roscardio.ru/ru/information-for-authors-ctp.html) Редакционные требования содержат 2718 слов. Описание статистических аспектов состоит из следующих 20 слов: «В таблицах должны быть чётко указаны размерность показателей и форма представления данных (M ± m; M ± SD; Me; Mo; перцентили и т. д.)».

11. Евразийский кардиологический журнал. (http://cardioweb.ru/files/Eurasian_Journal/1-2013.pdf) Редакционные требования содержат 1150 слов. В двух предложениях упоминаются статистические методы: «Приводятся количественные и качественные характеристики больных (обследованных), а также **упоминаются все методы исследований, применявшихся в работе, включая методы статистической обработки данных**. … **Указывайте статистические методы**, использованные для представления вариабельности данных и **достоверности различий**».

12. Журнал "Фарматека" (http://www.pharmateca.ru/ru/page/rules.html) Редакционные требования содержат 297 слов. Нет ни одного слова об описании статистики.

13. "Флебология". (http://www.mediasphera.ru/journals/flebo/pravila/) Правила для авторов содержат 1280 слов, в т. ч.:

«20. В оригинальных статьях необходимо указать, в каком из этапов создания статьи принимал участие каждый из ее авторов:
- Концепция и дизайн исследования

- Сбор и обработка материала
- **Статистическая обработка данных**
- Написание текста
- Редактирование

Пример:
Авт. И.И. Иванов, П.П. Петров, С.С. Сидоров
Участие авторов:
Концепция и дизайн исследования — И.И., С.С.
Сбор и обработка материала — П.П.
Статистическая обработка — П.П.
Написание текста — С.С.
Редактирование — И.И.»

Т.е. редакция журнала «Флебология» предлагает идентифицировать исполнителя статистической обработки данных, но при этом игнорирует сами требования к описанию результатов такой обработки.

14. "Российский кардиологический журнал" (http://www.roscardio.ru/ru/information-for-authors-rjc.html)
Редакционные требования содержат 3737 слов. Описание статистики содержит 57 слов:
«Статистика
Все публикуемые материалы рассматриваются экспертом по статистике и должны соответствовать "Единым требованиям для рукописей, подаваемых в биомедицинские журналы" (Uniform Requirements for Manuscripts Submitted to Biomedical Journals, Ann Intern Med 1997, 126: 36–47). **В подготовке статистической части работы рекомендуется использовать специальные руководства**, например, Европейского кардиологического журнала:
www.oxfordjournals.org/our_journals/eurheartj/for_authors/stat_guide.html **Статистические методы подробно описываются в разделе "Материал и методы"**».
К сожалению, требование подробного описания статистических методов редакция не конкретизировала, в т. ч. с не привела примеров такого описания.

15. Журнал "Рациональная Фармакотерапия в Кардиологии" (http://www.rpcardio.ru/author.php)
Редакционные требования содержат 3148 слов, в т. ч. **описание статистики содержит 168 слов**:

«7.5. Статистика
7.5.1. Все публикуемые материалы могут быть рассмотрены на соответствие и точность статистических методов, и статистическую интерпретацию результатов. В разделе Методы **должен присутствовать подраздел подробного описания статистических методов**, включая конкретные методы, используемые для обобщения данных; **методов, используемых для проверки гипотез** (если таковые имеются), и **уровень значимости** для проверки гипотез. При использовании более сложных статистических методов (помимо t-теста, хи-квадрата, простой линейной регрессии) должен быть указан статистический пакет, применявшийся при обработке результатов, и номер его версии.
7.9. Таблицы
7.9.1. Таблицы должны быть напечатаны через двойной интервал, иметь номер, соответствующий порядку упоминания в тексте, и название. Таблицы должны быть построены сжато, наглядно, иметь заголовки колонок и строк, строго соответствующие их содержанию. Данные, представленные в таблицах, не должны дублироваться в тексте или рисунке. В таблицах должны быть четко указаны размерность показателей **и форма представления данных (M ± m; M ± SD; Me; Mo; перцентили и т.д.)**. Все цифры, итоги и проценты в таблицах должны быть тщательно выверены, а также соответствовать своему упоминанию в тексте. При необходимости приводятся пояснительные примечания ниже таблицы».

Таким образом, можно констатировать, что из 15 проанализированных журналов кардиологической тематики лишь в журнале "Рациональная Фармакотерапия в Кардиологии" ясно

и чётко сформулированы конкретные требования к описанию статистических аспектов публикуемых результатов исследований. Однако отметим, что упоминаемое выше **руководство по описанию статистики от Европейского кардиологического журнала** (www.oxfordjournals.org/our_journals/eurheartj/for_authors/stat_guide.html) **является более подробным, и содержит 1976 слов**.

Вполне уместно задаться вопросом, почему же в самом журнале "Рациональная Фармакотерапия в Кардиологии", как и во многих других журналах, не сочли необходимым внести в редакционные требования перевод данного руководства, или собственную версию требований, основанную на его содержании? Быть может потому, что члены редколлегий этих журналов не публикуют собственные статьи в Европейском кардиологическом журнале, и потому не сталкиваются с проблемой возврата рукописей из-за низкого качества описания статистики? Или же вообще не читают ни этот журнал, ни упомянутое руководство по описанию статистики в публикациях? Впрочем, более вероятно, что основной причиной этого является совершенно иной мотив, который будет сформулирован ниже.

Анализ статей, публикуемых в описанных выше журналах, показывает, что уровень описания статистических аспектов в них, как и в редакционных требованиях, имеет прямую корреляцию с составами редакционных коллегий. Чем более корректно эти аспекты отражаются в публикациях членов редакционных коллегий и в редакционных требованиях, тем более полно эти аспекты отражены и в остальных публикациях данных журналов. Однако по нашему мнению гораздо большее влияние на содержание редакционных требований оказывает вот какой аспект. Многие журналы издаются частными издательствами. В других журналах за публикацию статьи с авторов взимают плату. Очевидно, что повышение уровня редакционных требований к описанию статистических аспектов проведённых исследований, уменьшит число авторов, желающих публиковаться в таких журналах. Поскольку для этих авторов проще будет опубликовать свои рукописи в тех журналах, в которых такие требования будут отсутствовать. Ясно, что снижение числа публикаций в журналах уменьшит доход у владельцев журналов.

Не менее важен и такой аспект, как доступность частных журналов для публикаций статей, которые пишутся «под заказ» фирмами, торгующими заказными диссертациями. Наконец, игнорирование положения с редакционными требованиями журналов со стороны таких организаций, как РАН, РАМН, Минздрав, Минобранауки и ВАК, также способствует развитию многих журналов как региональных «карманных изданий». В качестве примера приведём ситуацию с медицинскими журналами Кузбасса, рассмотренную в нашем обзоре «Доказательная или сомнительная? Медицинская наука Кузбасса: статистические аспекты» (http://www.biometrica.tomsk.ru/kuzbass1.htm) . После публикации данного обзора «Общество специалистов доказательной медицины» обратилось в ВАК с предложением об исключении этих журналов из перечня. Однако ВАК проигнорировало это обращение ОСДМ. И эту реакцию ВАК заранее предсказал частный владелец этих журналов, написав нам в своём письме: «Вы что, считаете, что ВАК подобно унтер-офицерской вдове сама себя высечет?» Проигнорировала аналогичное обращение и Е.К. Нечаева — директор Департамента научных и научно-педагогических кадров Министерства образования и науки РФ. Всё это наводит на мысль о

личной заинтересованности, а возможно, и личном участии в этом бизнесе многих из тех, от кого зависит уровень требований к диссертациям и научным публикациям [67].

Если расширить перечень анализируемых отечественных журналов, то можно найти единичные издания, в редакционных требованиях которых достаточно конкретно и ясно формулируются требования по описанию статистики. В качестве примера приведём ниже редакционные требования журнала «Экология человека», издаваемого в Южно-Уральском государственном медицинском университете (http://www.chelsma.ru/nodes/8438/), **которые содержат 3572 слова**. Ниже приведены фрагменты этих требований, имеющие непосредственное отношение к описанию статистических аспектов исследований.

«Данные требования составлены на основании «Единых требований к рукописям, представляемым в биомедицинские журналы», разработанных Международным комитетом редакторов медицинских журналов (ICMJE), в который входят редакторы таких всемирно признанных журналов, как «Журнал Американской медицинской ассоциации» (Journal of the American Medical Association, JAMA). Редакция благодарит Национальный институт общественного здоровья (г. Осло, Норвегия) и лично старшего советника института А. М. Гржибовского (В.Л.: см. [18]) за содействие в разработке данных требований.
…

Все статьи, поступившие в редакцию, подвергаются тщательному рецензированию. Рукопись, содержащая статистические данные, направляется помимо рецензента по соответствующей рубрике и рецензенту по статистике. Если у рецензентов возникают вопросы, статья возвращается авторам на доработку. Редакция имеет право запросить исходную базу данных, на основании которой производились расчёты в случаях, когда возникают вопросы о качестве статистической обработки.

Данные требования составлены с учётом основных недостатков отечественных исследований, описанных как в отечественной, так и зарубежной литературе. Особое внимание уделяется требованиям к описанию методологии проведения исследований, выбору методов статистической обработки данных, представлению результатов и их интерпретации. Редакция журнала надеется, что строгое соблюдение этих требований авторами рукописей поможет существенно повысить качество журнала и его цитируемость отечественными и зарубежными исследователями.

Методы

Раздел должен включать только те методы, которые предполагалось использовать на стадии планирования проекта согласно оригинальному протоколу исследования. Дополнительные методы, необходимость применения которых возникла в ходе выполнения исследования, должны представляться в разделе «Обсуждение результатов». Раздел «Методы» должен быть написан настолько подробно, чтобы читатель мог не только самостоятельно оценить методологические плюсы и минусы данного исследования, но при желании и воспроизвести его. Следует помнить, что опубликованное оригинальное исследование является не только источником информации, но и обучающим материалом для начинающих исследователей. Кроме того, подробное написание этого раздела позволит читателю самостоятельно оценить достоинства и недостатки исследования, а также достоверность и обобщаемость результатов.

В разделе рекомендуется представлять чёткое описание следующих моментов (выделение их в рукописи в отдельные подразделы необязательно): тип исследования; способ отбора участников исследования; методика проведения измерений; способы представления и обработки данных; этические принципы.

Тип исследования

В данном подразделе чётко обозначается тип проводимого исследования (обзор литературы, обсервационное, экспериментальное, и т. д.). При проведении обсервационного исследования следует указать, является ли оно описательным или аналитическим. В аналитических определяется разновидность исследования: поперечное (одномоментное), типа «случай – контроль», когортное, экологическое (корреляционное), и т. д. Рекомендуется указывать год и месяц проведения исследования, особенно в случае изучения признаков, которым характерна сезонная изменчивость. В литературных обзорах следует чётко указать, по каким критериям отбирались публикации для них.

Способ отбора участников исследования

В этом подразделе чётко указывается, каким образом отбирались пациенты или лабораторные животные для наблюдений и экспериментов. Обозначаются критерии для включения потенциальных участников в исследование и исключения из него. Рекомендуется указывать генеральную совокупность, из которой производится отбор участников исследования и на которую полученные результаты будут экстраполироваться. При использовании в исследовании такой переменной, как расовая или этническая принадлежность, следует объяснить, как эта переменная оценивалась и какое значение несёт использование данной переменной. В обсервационных исследованиях следует указывать способ создания выборки (простой случайный, стратифицированный, систематический, кластерный, многоступенчатый, и т. д.) и аргументировать включение в исследование именно этого количества участников. В экспериментальных указывать на наличие или отсутствие рандомизации участников исследования. Необходимо представлять описание процедуры рандомизации. Кроме того, следует указывать, проводилась ли процедура маскирования. Приветствуются расчёты минимального необходимого объёма выборки для проверки статистических гипотез или ретроспективный расчет статистической мощности для основных расчетов.

Методика проведения измерений

Все процедуры измерения тех или иных параметров, сбора данных, проведения лечебных или диагностических вмешательств должны быть описаны настолько детально, чтобы исследование можно было воспроизвести по представленному описанию. При необходимости можно сделать ссылку на детальное описание используемого метода. Если исследователь использует собственную модификацию ранее описанного метода или предлагает новый, то обязательно представляется краткое описание используемой модификации или предлагаемого метода, а также аргумент против использования общепринятых методов. Указываются названия лекарственных средств (как коммерческие, так и международные), химических веществ, дозы и способы введения препарата, применяемого в данном исследовании. Используемые аппараты, инструменты, лекарственные препараты и т. д. сопровождаются ссылкой на производителя.

Способы представления и обработки данных

Данный подраздел часто является основной причиной для отказа в публикации работ российских ученых за рубежом.
Описывать используемые методы обработки данных необходимо настолько подробно, чтобы читатель, имеющий доступ к исходным данным, мог проверить полученные результаты. Редакция журнала может в сомнительных случаях запросить у авторов статьи исходные данные для проверки представляемых результатов. В этом подразделе следует дать определение всем статистическим терминам, символам и сокращениям, используемым в работе. Например, M – среднее арифметическое, SD – стандартное отклонение, m – стандартная ошибка среднего арифметического, Me – медиана, Mo – мода, и т. д. Если в исследовании проверяются статистические гипотезы, то следует указывать принятый авторами критический уровень значимости (p). Гипотезы должны формулироваться чётко и описываться понятным читателю языком.

Редакция журнала не рекомендует полагаться исключительно на использование достигнутого уровня значимости при проверке статистических гипотез, так как величина p не отражает всей полноты информации.

Прочитай и перешли своим коллегам…

Рекомендуется представлять результаты с соответствующими показателями ошибок и неопределенности (доверительные интервалы). При описании статистических методов должны приводиться ссылки на руководства и справочники с обязательным указанием страниц. Помимо статистических процедур для проверки гипотез рекомендуется рассчитывать величину эффекта для наиболее важных сравнений.

Если в исследовании применяется несколько статистических критериев, следует упомянуть их все и указать, в какой ситуации каждый из критериев использовался. **Расплывчатое описание статистической обработки данных типа «вариационно-статистическую обработку проводили с помощью общепринятых параметрических и непараметрических методов статистики с использованием пакета прикладных программ Statistica» является неинформативным и недопустимым**. Работы с подобными формулировками будут сразу же отправляться на доработку без дальнейшего рецензирования, что значительно увеличит время от предоставления рукописи в редакцию до публикации. Применение тех или иных методов обработки данных должно чётко аргументироваться. Например, используя параметрические критерии, следует описывать, с помощью каких критериев проводилась процедура проверки распределения. Необходимо указывать, как производилась проверка соблюдения условий применения методов, для которых эти условия необходимы. Каждый из применяемых критериев должен быть обозначен так, чтобы исключить варианты прочтения. Например, если сравнение выборочных средних проводилось с помощью критерия Стьюдента, то следует указывать, какой из критериев Стьюдента (для независимых выборок или для парных наблюдений) использовался в работе. Недостаточно сказать, что применялся корреляционный анализ, надо указать, какой из коэффициентов корреляции рассчитывался. При использовании многомерных методов обработки данных указывается, каким способом отбирались переменные для включения в модели и какие категории использовались в качестве категорий сравнения. Если применяется редко встречающийся метод обработки данных, надо указывать, почему был выбран именно этот метод, представить ссылку на литературный источник и кратко описать используемый метод.

Если для обработки данных применяется пакет статистических программ, следует указывать его название, версию и производителя. Сообщать, на каком компьютере производилась обработка данных, ввиду отсутствия практической ценности данной информации, не нужно.
……
Результаты

Раздел предназначен только для представления основных результатов исследования. В этом разделе результаты, полученные в ходе данного исследования, не сравниваются с результатами аналогичных исследований других авторов и не обсуждаются.

Результаты следует представлять в тексте, таблицах и рисунках в логической последовательности исходя из очередности целей и задач исследования. Не рекомендуется дублировать в тексте результаты, представленные в таблицах или на рисунках, и наоборот. Описываются, выделяются и суммируются только важные наблюдения, относящиеся к задачам исследования. Общие характеристики выборки или исследуемых групп следует представлять в таблице с указанием основных изучаемых признаков. Необходимо указывать не только средние величины, но и меры рассеяния или доверительные интервалы для средних и долей.

Средние величины не следует приводить точнее, чем на один десятичный знак по сравнению с исходными данными. При описании долей необходимо указывать абсолютное количество наблюдений, особенно при малых выборках. Проценты приводятся с двумя десятичными знаками, только если доля составляет менее 1 %. Если доля составляет от 1 до 10 %, достаточно одного десятичного знака. Приветствуется использование 95 % доверительных интервалов, как для средних величин, так и для долей.
Достигнутый уровень значимости для каждого из использованных статистических критериев следует представлять с точностью до трех десятичных знаков. Редакция настоятельно рекомендует избегать

формулировок типа p < 0,05 (исключение составляет ситуация, когда p < 0,001). Вместо формулировок p > 0,05 или «различия незначимы» следует указывать абсолютное значение величины p. Помимо достигнутого уровня значимости рекомендуется приводить фактические значения критериев и число степеней свободы. Например, критерий хи-квадрат Пирсона при наличии двух степеней свободы представляет собой следующее: $\chi 2 = 29,2$, d. f. = 2, $p < 0,001$.
Всегда следует помнить, что выявление статистически значимых различий ещё не означает наличия достоверных или клинически важных различий, а также причинно-следственных связей».

Итак, можно констатировать, что, несмотря на крайне низкий уровень требований к описанию статистических аспектов в большинстве журналов, всё же имеются прецеденты весьма качественных описаний требований статистических аспектов в некоторых журналах. К сожалению, эти прецеденты не становятся для редакций кардиологических журналов примерами для подражания. Причины отсутствия такого подражания могут быть различны. В том числе весьма вероятна следующая причина. При сравнении авторских составов проанализированных статей 2-го и 3-го разделов, мы провели частотный анализ этого состава. И обнаружили, что подчас одни и те же авторы встречаются в обоих разделах. Такая ситуация наводит на мысли о том, что у разных авторов и редколлегий, могут быть разные представления о целях публикаций. Низкий уровень качества публикаций может быть проявлением мнения о невысокой ценности и важности публикаций для повышения качества здравоохранения. Например, хирурги перед проведением операций проводят тщательную коллективную подготовку к ним. Такая тщательность является обязательным и необходимым условием положительного исхода операции. Тогда как такая корреляция между качеством публикуемых в журналах статей и уровнем здравоохранения признаётся не всеми.

Данные + Статистика = Знания + Экономика

Ключом ко всякой науке
является вопросительный знак

Оноре де Бальзак

Как и в статье «Применение методов статистики в кардиологии (по материалам журнала «Кардиология» за 1993–1995 гг.) опубликованной 15 лет назад, в нынешнем обзоре также констатируем наличие такого «заболевания», как «патологическая статистика» в кардиологических публикациях. Причин длительного существования такой журнальной патологии немало. Часть из них рассмотрены выше, другие подробно на сайте БИОМЕТРИКА (http://www.biometrica.tomsk.ru/), которому в ноябре 2013 года также исполняется 15 лет. Однако если провести интеграцию этих причин, то в итоге можно сформулировать доминирующий аспект, который, в конечно счёте, и объясняет появление этой патологии.

Весьма важно в данном случае ответить и на такой вопрос: является ли данная патология присуща только отечественной кардиологии, или же она имеет более широкий характер? О

системном характере этого явления говорит тот факт, что оно наблюдается не только в кардиологии, и не только в наше время, но и в целом в отечественной медицине и биологии, и в гораздо более раннем периоде. В нашей статье «Долгое прощание с лысенковщиной» [29], написанной 20 лет назад, мы приводим пример того, как Н.И. Ермолаева, аспирантка академика Т.Д. Лысенко, опубликовала статью, в которой с помощью статистики пыталась опровергнуть законы Менделя. И поскольку она была уверена в истинности своих результатов, то она опубликовала даже таблицу проанализированных ею данных. Что и позволило затем академику А.Н. Колмогорову, проведя корректный анализ этих данных, опубликовать в «Докладах Академии наук СССР» свою статью «Об одном новом подтверждении законов Менделя.

О системном характере статистической некорректности в медицине говорит и книга профессора Института организации медицины А.Я. Боярского «Статистические методы в экспериментальных медицинских исследованиях», опубликованная в 1955 году. Автор детально проанализировал большое количество статей, опубликованных в течение нескольких лет в известных советских медицинских журналах разной тематики. По сути дела это было первое отечественное исследование, в котором автор на значительном конкретном материале констатировал крайне неблагополучную ситуацию с применением статистики в экспериментальной медицине и биологии тех лет. «Уже беглое ознакомление с состоянием дела показывает, что статистическая обработка экспериментальных данных является наиболее слабым местом во многих исследованиях. … **Трудно требовать от медика, чтобы он**, наряду со знаниями в своей собственной области, **был в то же время достаточно вооружённым**, скажем, в радиотехнике для конструирования аппаратуры, улавливающей биотоки, или **в статистике для нахождения наиболее правильных методов статистической обработки своих экспериментальных данных**. И подобно тому, как медику, несомненно, приходится обращаться за содействием к радиотехнику, для правильной статистической обработки экспериментальных данных нередко приходится обращаться к специалисту-статистику … Так или иначе, но **бесспорным фактом являются и недостаточная вооружённость медиков статистическими знаниями, и недостаточно высокий научный уровень статистической методики в большинстве их экспериментальных работ**» [7]. Напомним, что **эти слова были сказаны еще в 1955 г**., т.е. уже **более полувека назад констатировалось, что медицинские исследования по своей сути являются междисциплинарными, и уже потому требуют обязательного участия биостатистика.**

Аналогичный обзор «О применении статистических методов в медико-биологических исследованиях» был опубликован в 1987 г. профессором А.И. Орловым в журнале «Вестник Академии медицинских наук СССР» [49]. «… можно сделать вывод о целесообразности организации работ по критическому анализу сложившейся в медико-биологических исследованиях практики статистической обработки данных. … Целесообразно создание службы статистических консультаций в системе научно-исследовательских учреждений медико-биологического профиля».

«В известной среди медиков и биологов книге «Медико-биологическая статистика», её автор, английский медик С. Гланц, так пишет о патологической статистике. «Врачам известно

множество методов диагностики и лечения, эффективность которых **была «доказана» статистическими методами** и которые, тем не менее, канули в Лету, не выдержав испытания практикой. ... **Вред, приносимый ошибками такого рода, очевиден. Исследователь заявляет о «статистически достоверном» эффекте лечения, редактор помещает статью в журнал, врач, неспособный критически оценить публикацию, применяет неэффективный метод лечения. В конце этой цепи находится больной, который расплачивается за всё, подвергаясь ненужному риску и не получая действительно эффективного лечения**. Не следует сбрасывать со счетов и ущерб от самого факта проведения бессмысленных исследований. Деньги и подопытные животные приносятся в жертву науке, больные рискуют ради сбора ошибочно интерпретируемых данных» [34].

А вот как гораздо раньше об этом же сказал известный советский хирург-кардиолог, директор Института сердечно-сосудистой хирургии Николай Михайлович Амосов (для справки: в 1939 году Н.М. Амосов получил диплом медика с отличием, а в 1940 году – диплом инженера с отличием; в Институте сердечно-сосудистой хирургии он сделал 56 тысяч операций на сердце). «Медицина знает много подобных историй. Новые лекарства, операции. Потом оказывается – бесполезные, а то и вредные. Без этого нельзя. Не всё можно проверить на собаках и мышах. Только через грустный опыт отстаивается золотой фонд медицины. Врачи тоже люди, и они искренне увлекаются. Материал для экспериментов, к сожалению, неподходящий. Но другого нет**. Одно нужно – честность. Признайся, удержи других. Но на это тоже нужно мужество**. Ох, как не хочется признаваться! … «Право учёного на ошибки». Нельзя его отнять даже в медицине. Не рисковать – значит замедлять темп развития, масса людей погибнет, не дождётся. Человечеству обойдётся много дороже» [3].

«Наш многолетний опыт общения с медиками и биологами показывает, что на этом этапе главным препятствием для последующего результативного статистического анализа собранных экспериментальных данных является их идейная (в смысле идей статистики) неподготовленность. Эта идейная безграмотность, проявляемая в незнании основных возможностей современной прикладной статистики, приводит к примитивизму выводов, получаемых после анализа нередко бесценных наблюдений. Однако вернемся к этапу декомпозиции целей исследования в конкретные задачи, решение которых и должно обеспечить исследователю достижения поставленной цели. Часть сформулированных задач для их решения требуют статистического аппарата. Ниже мы попытаемся проиллюстрировать этот этап, используя вполне конкретный пример из нашей практики. Отметим, что данный этап декомпозиции целей исследования в задачи, требующие последующего статистического анализа данных, лучше всего выполнять совместно с квалифицированным биостатистиком. Более оптимально, если этот же специалист будет сам организовывать или решать сформулированные задачи статистического анализа данных. … опыт разработки задач анализа биомедицинских данных позволяет выделить два подхода в реализации этого этапа.

Первый подход характерен для небольших исследований, имеющих локальный характер. Например, при написания статьи или доклада на конференцию или в журнал, либо же на этапе написания дипломной работы требуется провести анализ небольшой выборки наблюдений,

порядка 20-40 наблюдений с 4-8 признаками, переменными. Обычно такие работы требуют для своего выполнения 2-3 вполне определенных и конкретных методов, о которых сам исследователь, как правило, может уже достаточно внятно объяснить. К примеру, требуется сравнить средние значения конкретного количественного признака в группе больных и в контрольной группе здоровых. Либо же провести аналогичное сравнение группы больных до лечения, и после лечения и т.д. При реализации такого подхода в формулировке задач и выборе необходимых методов их решения доминирует сам исследователь. На долю биостатистика в этом случае остаётся либо роль квалифицированного и оперативного исполнителя анализа данных, либо же, исходя из особенностей признаков, некоторая корректировка предлагаемых методов анализа.

Второй подход характерен для исследований большего объёма – НИОКР, кандидатские и докторские диссертации и т.д. Необходимым условием реализации такого подхода является наличие истинного профессионализма у руководителя исследования. Такой профессионал не только осознает ограниченность собственных знаний, но и не стыдится этой ограниченности. К сожалению, такое осознание приходит с годами и не свойственно молодым начинающим исследователям. Подобный профессионализм приводит к пониманию того, что только участие такого же профессионала в области анализа данных позволит максимально полно использовать имеющиеся наблюдения и получить адекватные результаты» [31].

Проблема методологического конфликта в науке существует давно. Так 125 лет назад, 1 сентября 1888 года, вступительной лекцией профессора С.И. Коржинского на тему «Что такое жизнь?» была открыта подготовка медиков в Томском Императорском университете – первом за Уралом университете, основанном в 1878г. В своей лекции С.И. Коржинский коснулся актуальной для нас и сегодня темы методологии науки [31].

«Таково уж свойство человеческого ума, что как только удовлетворены все насущные житейские потребности, человек стремится постигнуть смысл окружающих явлений, подметить законы, по которым они совершаются. Он делает наблюдения, собирает факты, выводит заключения; но чем глубже проблема, тем слабее, ничтожнее кажутся фактические знания, которыми обладает человечество, и страстная потребность ума найти решение волнующих его вопросов удовлетворяется иным путем - тем полетом мысли, который называют воображением. Таким образом, два орудия мысли находятся у человечества - факт и гипотеза; два направления - эмпиризм и философия… Факты есть основа науки; они лежат в основе и всякой гипотезы. Факты - это золото, которое тщательно сосчитанное и рассортированное лежит в кладовых банка. Теории и гипотезы - это кредитные бумаги, которые пускаются в обращение, и которые имеют цену, сообразную с количеством звонкой монеты, которому она соответствует. Факты имеют значение абсолютное, теории и гипотезы - лишь относительное. **Однако одни Факты ещё не составляют науки.** Грубый эмпиризм не может удовлетворить запросов ума. **Работа учёного не может состоять лишь из собирания фактов. Для него необходимо** возвышаться над фактами, осматривать их с высоты птичьего полета, стремиться **уловить внутреннюю связь явлений**, восполняя пробелы воображением, намечая пути новых исследований. Если рассматривать вблизи какую-нибудь картину, то мы увидим лишь мазки кисти и грубые очертания. Чтобы уловить идею картины, чтобы постичь ее красоту, нужно отойти на

некоторое расстояние. **Факты есть строительный материал**, теория - план будущего здания. **Во время постройки план может много раз меняться, строительный материал идёт в дело, так или иначе, но без плана невозможно вывести здание. Так без теории не может существовать наука**».

Эволюционизм науки, о котором говорил С.И. Коржинский, есть обязательное условие развития науки. При этом эффективное развитие науки реализуется через обучение и творчество. Эффект от реализации двух этих компонент науки возможен лишь при паритете между поступлением новой информации, и её переработкой. Лишь в этом случае новые знания смогут быть усвоены и упорядочены. А это требует самостоятельных усилий по осмыслению каждой новой порции фактов, и полученных на их основе знаний. Отсутствие такой гармонизации новых и старых знаний только дезорганизует мышление. Вот как высказался об этом один из родоначальников эволюционизма, английский философ Герберт Спенсер: "**если знания человека находятся в беспорядочном состоянии, то чем больше он имеет их, тем сильнее расстраивается его мышление**". Фактически в данном случае речь идёт о необходимости оптимальных пропорций между широтой и глубиной знаний у конкретного исследователя. Чем шире область интересов исследователя, тем меньше глубина знания. И наоборот. О важности этого аспекта вообще в жизни можно судить, например, вот по какому примеру. Если сделать в поисковой системе Яндекс запросы типа «оптимальное соотношение ширины и длины», «оптимальное отношение глубины и ширины», и т.п., то получим примерно 400-600 тысяч ссылок. Причём в некоторых из этих запросов на первом месте среди найденных ссылок будут расположены весьма необычные, и даже неожиданные материалы…

Весь предыдущий период развития человечества характеризовался изменением приоритетности между такими основными ресурсами, как:

- материальные ресурсы
- природные ресурсы
- трудовые ресурсы
- финансовые ресурсы
- энергетические ресурсы.

Однако начиная со второй половины 20-го века достаточно большую долю в развитии общества стали занимать информационные ресурсы. Выход в 1984 году книги Громова Г.Р. «Национальные информационные ресурсы: проблемы промышленной эксплуатации» [19], пояснил, почему данные ресурсы стали рассматриваться в мире как одна из экономических компонент. Это объяснялось растущей зависимостью экономик развитых стран, как от объёмов имеющейся информации, так и от эффективности средств передачи и переработки информации. Занимавшийся этой проблемой профессор Гарвардского университета А. Осттингер писал: «…наступает время, когда **информация становится таким же основным ресурсом, как материалы и энергия**, и, следовательно, по отношению к этому ресурсу должны быть сформулированы те же критические вопросы: кто им владеет, кто в нём заинтересован, насколько он доступен и возможно ли его коммерческое использование».

Прочитай и перешли своим коллегам…

О том, что роль информации как экономической категории станет ведущей, ещё в 1967 году писал Р. Сарноф, президент крупнейшей радиокорпорации США: «...недалеко то время, когда информация станет товаром первой необходимости, равным по важности энергии». Сейчас информационные услуги обеспечивает в США более 50 % доходов национального бюджета страны, а специальные службы США на сбор и обработку информации расходуют более 2/3 своего бюджета.

Проецируя данный аспект развития цивилизации на медицину, возникают аналогичные вопросы о соотношении приоритетов между всеми ресурсами. При этом очевидно, что эти соотношения будут различаться в медицинской практике и науке. Точно также можно говорить и о различной ценности информации в двух этих областях медицины. Прежде чем попытаться сформулировать основную причину живучести обсуждаемой статистической патологии, зададимся вопросом: – А вообще-то нужна ли кардиологии статистика? ***А если нужна, то зачем и почему***? На первый взгляд, этот вопрос покажется примитивным и не очень уместным для данного обзора. Вроде бы и так всё ясно и понятно: да, кардиологам статистика нужна. Кто же в этом сомневается и оспаривает эту необходимость? Однако в действительности этот вопрос не столь прост, как может показаться на первый взгляд. Ниже мы обсудим этот аспект, и попытаемся более подробно ответить на него. Для того чтобы результат такого обсуждения был более доступным для понимания именно медиками, воспользуемся для этого публикациями самих учёных-медиков и преподавателей медицинских вузов. В частности, обсудим фрагменты из популярных учебников, монографий и статей.

Учитывая межпредметный, философский характер этого вопроса, обратимся к книге «Философия науки и медицины» [63], автор которой известный специалист в области истории и философии медицины и науки, профессор, заведующий кафедрой философии Первого Московского государственного медицинского университета (ПМГМУ) им. И.М.Сеченова, Хрусталёв Ю.М.

«Что касается развития научной медицины в античной Греции, то это происходило противоречиво. Согласно Гиппократу, она долгое время оставалась **врачебным искусством.** Медицина не могла стать научной, пока не произошло её соединение с натурфилософией. … Но по-настоящему научно-философский подвиг в медицине совершил великий врач и мыслитель античности Гиппократ (460-377 до н.э.). Он, по сути, выделил из натурфилософии медицину как самостоятельную дисциплину. Учёный первым в истории медицины **придал ей и статус науки,** что означало создание принципиально нового рода занятий врачей - **исследования человека и причин его заболевания**. Они стали активно использовать **точные (научные) методы** опознания болезней. Согласно Гиппократу, медицина делает всех врачей истинными философами, т.е. обеспокоенными судьбой людей. Это – благородное служение людям. Оно даёт врачам ощущение высокой гуманитарной ответственности за качественное предупреждение или лечение заболеваний. А посему *философская мудрость нацеливает медика на* ***постоянное стремление к овладению научными знаниями, передовыми методами,*** *лучшими средствами избавления людей от болей*: больших и малых, физических и душевных. Поэтому медику, как мудрому человеку, укрепляющему здоровый дух

Прочитай и перешли своим коллегам…

людей, требуется глубокое специальное и социально-культурное ***осознание медицины как естественно-научной и гуманитарной науки***… Последователь врачебной идеологии Гиппократа Гален (ок. 130 - ок. 201) в «Истории философии», написанной для студентов-медиков, не уставал повторять, «что лучший медик должен быть истинным философом». Медикам своего времени, отвернувшимся от учения и морали Гиппократа, Гален предъявил три обвинения: в невежестве, коррупции и абсурдной разобщённости. **В первом – он говорил об устранении врачей от научного познания человека. Утратив его, медицина перестает быть врачеванием и превращается в «ползучую эмпирическую практику»** [63].

Таким образом, в принципе кардиолог может **эмпирически добывать это знание на собственных пациентах**. Очевидно, что цена такого знания будет катастрофически высокой и неприемлемой для общества. В таком случае лечащий врач-кардиолог, не являясь учёным, исследователем, однако желая «овладевать новыми научными знаниями» в области кардиологии, будет вынужден читать статьи в кардиологических журналах, диссертации по кардиологии, слушать доклады на конференциях, и т.д. Но тогда ему необходимо иметь минимальный набор знаний по статистике, которую используют авторы публикаций, и докладчики, обобщая собственный эмпирический опыт методами статистики. Это позволит ему оценивать надёжность и корректность авторских выводов, приведённых в этих публикациях, и самостоятельно принимать решение о целесообразности их использования в своей врачебной деятельности. Т.е. избегать тех ситуаций, о которых писали Н. Амосов и С. Гланц.

Именно этот аспект, определяющий уровень обучения статистике лечащих врачей, а не учёных-медиков, и был подробно описан в нашем докладе [46] на Международной конференции по доказательной медицине в Ереване (18-20.10.2012). Для лечащего врача, а не для учёного-медика, исследователя, объём и уровень познаний в области статистики определяется главной целью использования этих знаний: определить надёжность авторских выводов в публикациях, и сделать собственное заключение о доверии или недоверии к этим выводам. Чтобы далее принять или отвергнуть их в своей практической деятельности. Для учёных-медиков, исследователей, область применения статистических знаний совершенно иная. Конечно, эти знания, как и у практикующих врачей, нужны им для оценки публикаций. Такие публикации близки к тематике их исследований. Что уже сужает набор статистических методов, которыми необходимо владеть таким исследователям. Однако уровень владения такими методами должен быть таковым, что часть из них, не самых сложных, исследователь мог бы выполнять самостоятельно, без привлечения профессиональных биостатистиков. Этот аспект требует определённого уровня теоретических знаний по данным методам. Отсутствие же таких знаний как раз и приводит к тем отрицательным примерам, которые были приведены выше. Важным аспектом этого требования является то, что результат некачественного статистического анализа по своей сути является дезинформацией. Далее эта дезинформация, будучи применена во врачебной практике, вполне может оказать отрицательное воздействие на пациентов, что является явным нарушением гиппократовского принципа «Не навреди!».

Вторая компонента определения объёма и уровня знаний у медиков в области статистики определяется спектром статистических методов, используемых в публикациях, и уровнем их

описания. Например, если почитать статьи в медицинских журналах и диссертациях, написанные 50-60 лет назад, то можно обнаружить, что в них присутствует очень небольшой набор статистических методов. Ещё меньше этот набор методов в публикациях 100-летней давности. Однако ошибочно думать, что расширение набора используемых в медицине статистических методов обусловлено лишь развитием самой статистики. Причины этого, прежде всего, в развитии самой медицины. Вот что пишет об этом Хрусталёв Ю.М.

«В связи с возникновением новых пограничных дисциплин (медицинская радиология, медицинская генетика, клиническая физиология, космическая биология и медицина) всё большее значение приобретают такие философские и методологические аспекты, как **взаимосвязь физических, химических, биологических и социальных закономерностей в жизнедеятельности организма** (в норме и патологии), соотношение различных приёмов исследований, **возможность и сфера применимости математических, кибернетических, *статистических*** и других средств познания. Это обусловливает возможность применения новых научных методик изучения, оперирующих количественными величинами и их соотношениями» [63]. Т.е. для учёта многообразных и сложных количественных взаимосвязей и закономерностей, требуются адекватные, **столь же многообразные и сложные средства их познания, в т. ч. и сложные статистические методы.** И здесь также соблюдается общеизвестное соотношение между спросом и предложением. Например, многие из современных статистических методов были развиты как раз благодаря потребности в них, спросу на их использование. Если же исследователь вместо современных сложных методов статистического анализа зацикливается на самых примитивных методиках, например, лишь на сравнении средних, то этим он демонстрирует свой примитивизм восприятия сложнейших процессов происходящих в человеческом организме.

В развитии каждой науки есть свои этапы перехода в новый качественный уровень. Таким новым этапом в настоящее время является концепция доказательной медицины. Не случайно эта концепция зародилась именно на Западе, где спектр и уровень использования возможностей различных наук и техники, в том числе и статистики, в медицинской науке несравненно выше, чем в странах бывшего СССР. Доказательная медицина «…руководствуется двумя ключевыми принципами:

- каждое клиническое решение врача должно приниматься с учётом научных данных;
- **ценность каждого нового факта тем выше, чем основательнее выработана методика научного исследования.**

Именно они становятся важнейшей частью научной идеологии всей системы здравоохранения» [63].

Одним из показателей уровня научного исследования и публикаций, является адекватность использования авторами специальной терминологии. Поскольку **адекватное владение терминами соответствующей области знания отражает корректное усвоение авторами базовых понятий этой науки**. Обратимся вновь к монографии Хрусталёва Ю.М. «В научно-познавательной деятельности учёного большую роль играют термины (лат. terminus - граница,

предел). Под термином понимается слово или словосочетание, точно обозначающее понятие и его взаимосвязь и соотношение с другими понятиями. Характерной чертой терминов является автономность, отсутствие жёсткой связи с контекстом изложения того или иного научного сюжета. Термин должен быть максимально однозначным и нейтральным. … В условиях усиливающихся международных связей однозначное употребление научных терминов на разных языках – прогрессивное явление, облегчающее взаимопонимание ученых. **К сожалению, медицинская понятийная терминология не в полной мере отвечает выше перечисленным требованиям**» [63].

Для иллюстрации последнего утверждения рассмотрим систематически ошибочно используемый в публикациях по кардиологии термин «достоверность». Типичные обороты с этим термином таковы: «Различия считали достоверными», «Статистически достоверными считали различия», «достоверная разница», «достоверное отличие» и т.п. Как правило, такие обороты используются авторами при описании результатов проверки статистических гипотез с помощью тех или иных статистических критериев. Некорректность данного выражения уже давно описана известным специалистом в области доказательной медицины Н. Зориным в статье «**О неправильном употреблении термина "достоверность" в российских научных психиатрических и общемедицинских статьях**» [24] опубликованной нами на сайте БИОМЕТРИКА ещё в 2000 г. Поскольку подобные критические статьи не являются неким регламентом, которому обязательно должны следовать читатели, то обратимся к Государственным стандартам.

Как известно, практика создания государственных стандартов в России начинается с 1925 г. Конституция Российской Федерации относит такие стандарты к вопросам ведения государства. **Такие стандарты имеют обязательный характер, и их несоблюдение преследуется по закону.** Рассмотрим содержание Государственного стандарта ГОСТ Р 50779.10-2000 «**Статистические методы. Вероятность и основы статистики. Термины и определения**», М.: Госстандарт России. Этот стандарт состоит из 49 страниц, в которых на вполне доступном уровне объясняются базовые понятия и термины теории вероятностей и статистики. В предисловии читаем: «Разделы настоящего стандарта, за исключением разделов 1a, 1b и приложения А, **представляют собой аутентичный текст международного стандарта** ИСО 3534-1–93 «Статистика. Словарь и условные обозначения. Часть 1. Вероятность и основные статистические термины». Т.е. аналогичный международный стандарт был принят ещё 20 лет назад в 1993 году.

На 1-й странице данного ГОСТ'а читаем: «Настоящий стандарт устанавливает термины и определения понятий в области теории вероятностей и математической статистики». Отметим, что описания этих терминов вполне доступны для понимания и восприятия всем категориям читателей, и поэтому могут считаться определениями. Например, на стр. 2 в разделе 1 «Термины, используемые в теории вероятностей», читаем:

«*1.1 вероятность*

Действительное число в интервале от 0 до 1, относящееся к случайному событию». На этой же странице, справа от этого предложения, приведён перевод термина «вероятность» на английском и французском языках.

Напомню, что выше, в статье «Психогенная одышка и гипокапния у больных ишемической болезнью сердца до и после коронарного шунтирования» (ПКК, № 1, 2012, с. 39-42, было обнаружено предложение, в котором использовалось выражение «**p > 1**». Т.е. ни 5 авторов статьи, ни весь коллектив редакционной коллегии ПКК, допустивший рукопись этой статьи к публикации, и не подозревают о подобном ограничении на величину вероятности. Эту же самую информацию об интервале значений вероятности можно найти даже в следующих школьных учебниках:

1) Бунимович Е.А., Булычев В. А. Вероятность и статистика. 5-9 кл.: Пособие для общеобразоват. учеб. заведений. - М.: Дрофа, 2002. - 160 с.: ил.

2) Бородкина В.В. Высоцкий И.Р., Захаров П.И., Ященко И.В. Теория вероятностей и статистика. Контрольные работы и тренировочные задачи. 7-8 класс. – М.: МЦНМО, 2011. – 72 с.

3) Бродский Я.С. Статистика. Вероятность. Комбинаторика. М.: ООО «Издательство Оникс: ООО «Издательство «Мир и Образование», 2008. – 544 с.: ил. – (Школьный курс математики).

4) Мордкович А.Г. События. Вероятности. Статистическая обработка данных: Доп. Параграфы к курсу алгебры 7–9 кл. общеобразоват. учреждений/ А.Г. Мордкович, П.В. Семёнов. – 5- изд. – М.: Мнемозина, 2008. – 112 с.: ил., и т.д.

Если же внимательно прочитать весь ГОСТ «от корки до корки», то мы не обнаружим в нём ни одного выражения, использующего слово «достоверность». Зато найдём в нём много определений терминов, связанных с проверкой статистических гипотез. В частности, таких терминов, как «уровень значимости», «критическая область», «вероятность ошибки первого рода», «значимый результат», и т.д. Впрочем, для того чтобы познакомиться с этими определениями, вовсе не обязательно изучать данный ГОСТ. Вполне можно обойтись и учебниками математической статистики, однако написанными специалистами в этой области.

Обратимся вновь к монографии Хрусталёва Ю.М. «… исторически и логически научное понимание качеств вещи, предмета или явления предшествует познанию количественных отношений. Это объяснимо: прежде чем считать и измерять что-то, **человек должен понять, что же он подвергает счёту или измерению. … Количественные соотношения как более сложные познаются не сразу.** Качественное отличие и своеобразие цветов известно давно. По истечении длительного времени были раскрыты количественные особенности, объясняющие качественное различие цветов (длин волн). Познание количественного аспекта изучаемых систем и объектов свидетельствует о переходе на новый, более глубокий уровень познания. … С полным основанием можно сказать, что сейчас наступает качественно новый этап в изучении и практическом использовании количественных, математически выраженных закономерностей. **Но пока уровень количественной, математической зрелости медико-биологических наук явно недостаточен.** Даже при широком использовании

количественных характеристик (количество лейкоцитов, тромбоцитов) врачи имеют **дело не с полноценным количественным анализом** медико-биологических процессов, **а с активным использованием количественных характеристик и показателей** для качественного анализа, т.е. для описания и иллюстрации качественного состояния того или иного нормального или патологического процесса. Другими словами, количественные показатели используются **для характеристики** уже известных качественных различий и особенностей» [63].

Об истинности данного утверждения свидетельствует наличие в статьях многочисленных таблиц с выборочными **характеристиками** в виде средних, ошибок средних, стандартных отклонений и т.д. И гораздо меньшие по объёму разделы с результатами количественного **анализа взаимосвязей этих показателей**. А ведь именно модели таких многомерных взаимосвязей различных подсистем человеческого организма и позволяют выбирать оптимальные для медицинской практики комбинации лечебных воздействий. Именно учитывая эти взаимозависимости врачи и назначают комплексное лечения для конкретного пациента.

Продолжим чтение монографии Хрусталёва Ю.М. «В настоящее время происходит методологическая переориентация в понимании роли и значения количественных методов в изучении специфики качественной характеристики изучаемых медициной процессов. Широкое внедрение математических, кибернетических методов в медицину и здравоохранение должно строиться на использовании количественного и качественного анализа. … Возрастание интереса к проблеме измерения изучаемых медико-биологических явлений, их квантификации (лат. quantum - сколько), т.е. количественному измерению качественных признаков, сопровождается **усилением математизации и электронизации медико-биологических наук**. Иногда говорят даже о недостаточности классического математического аппарата для описания и объяснения сложных медико-биологических явлений и законов, лежащих в их основе, **так как в мире живой природы имеет место взаимодействие большого количества переменных величин, имеющих вероятностную природу**. Возникающие новые «метрии», «графии», «скопии» в лабораторном и клинико-диагностическом секторе медицины, как правило, опираются на широкое использование количественных, математических, электронно-кибернетических методов и устройств» [63]. Можно привести массу новых медицинских технологий, которые в течение разных периодов времени разрабатывались на основе этих методов, и сегодня используются в медицинской практике. Типичным примером такого устройства является магнито-резонансная томография, изобретатели которого в 2003 году получили Нобелевскую премию. Однако, несмотря на всю сложность этого метода, он не является универсальным методом, используемым для диагностики.

Вновь возвращаемся к монографии Хрусталёва Ю.М. «Формируются математическая биология и медицина, которые все более подводят к мысли, что в основе медико-биологических процессов лежит не просто количественно усложненные физико-химические закономерности, а иные, качественно отличающиеся законы. … Возрастание интереса медиков и биологов к проблеме измерения и расширение рамок его практического использования являются **свидетельством постоянного приближения медико-биологических наук к точным, математизированным наукам**. …Современная медицина уже не ограничивается познанием свойств, отражающих лишь качественную сторону изучаемых процессов и явлений. **Для**

углубленного познания ей необходимо выявить их количественную характеристику. Внедрение в медицинскую науку и практику измерительных средств и методов усилило объективную компоненту медицинского знания и умения. **Свойства, присущие медико-биологическим объектам, являются более сложными по сравнению со свойствами объектов, изучаемыми науками механического, физического или химического профиля**» [63].

Именно это последнее утверждение из монографии Хрусталёва Ю.М. и является главной причиной того, что медицина, в сравнении со всеми остальными науками, требует для своего развития привлечения гораздо большего количества других наук и технологий. Действительно, человек, будучи биосоциальным существом, является самым сложнейшим объектом для научных исследований. Однако расширение спектра используемых в медицине наук не может приводить к аналогичному расширению спектра знаний по всем этим наукам у каждого специалиста, причём с одинаковой глубиной познания. Выше мы уже писали о необходимости оптимальных пропорций между широтой и глубиной знаний у конкретного исследователя. И в силу этих ограничений для продуктивного использования новых наук и технологий исследовательским организациям необходимо вводить в свои структуры специалистов по этим новым направлениям.

Однако тот факт, что человек является самым сложнейшим объектом для научных исследований, не является единственной причиной, по которой отечественная медицинская наука весьма существенно отстаёт от передовых стран мира. В статье «Долгое прощание с лысенковщиной» [29], написанной нами 20 лет назад, мы приводим пример того, как ВАК СССР отказывала в присуждении докторских и кандидатских учёных степеней медикам, по причине использования ими «буржуазной статистики». За время гонений в СССР на генетику и кибернетику («продажных девок империализма», как тогда называли эти науки в СССР), медицину «отлучили» от статистики. И эта акция до сих пор сказывается на уровне использования статистики в медицинской науке. Таким образом, не менее важным аспектом, влияющим на уровень развития медицинской науки, является отношение к этому органов государственного управления. Можно привести массу примеров того, как государственные структуры, фактически не выполняя свои обязанности, тем самым снижают производительность и эффективность медицинской науки. Это в полной мере относится к кризисной ситуации с ВАК, отсутствию регулирования уровня качества журнальных публикаций путём разработки стандартных редакционных требований по различным отраслям науки, и т.д. Таким образом, специфичность объекта медицины, требующего привлечения большого спектра наук и технологий, как раз и объясняет трудности получения необходимой информации, и далее нового знания, порождающего новые, экономически более эффективные медицинские технологии. Повышение экономической эффективности медицинских технологий сопровождается при этом снижением энергетических и материальных затрат.

В настоящее время концепция триединства материи, энергии и информации становится доминирующей. Действительно, чтобы получить любое составляющее из этой троицы, необходимо обязательное использование остальных двух компонент. Проблема взаимосвязи этих трёх компонент весьма объёмна и сложна по своему содержанию, и её решение, как минимум,

заслуживает Нобелевской премии. Поэтому ниже приведём только краткие выводы, касающиеся этого аспекта.

Во-первых, получение всякой информации требует определённых объёмов энергии и материи. Простейший пример этому – «демон Максвелла» (если вы не знаете что такое «демон Максвелла», сделайте запрос в Яндексе, и найдёте десятки тысяч ссылок с пояснением этого выражения). Учитывая то, что этот пример может быть знаком не всем читателям данного обзора, приведём другой пример. Так, для получения томограммы требуется наличие как самого томографа (материальный ресурс), так и определённого количества электроэнергии (энергетический ресурс). Но помимо этих двух компонент необходимо также и наличие информационного ресурса, которым располагает специалист, работающий на томографе. А далее подобным информационным ресурсом должен владеть и лечащий врач, который изучает томограммы.

Второй аспект этого триединства – взаимозаменяемость этих ресурсов. Например, прежде рост объёма телефонных и телеграфных сообщений происходил за счёт увеличения протяжённости линий из меди. Однако развитие информационных ресурсов, приведших к созданию технологий на основе мобильной и оптоволоконной связи, резко сократило потребности в этом металле (материально-энергетический ресурс), и в то же время резко повысило объёмы и скорости передачи информационного ресурса. Аналогичная взаимозаменяемость есть и в лекарственных средствах. Анализ взаимозаменяемости ресурсов привёл к появлению в экономической теории термина «эластичности» ресурсов. Т.е. увеличение информации содержит в себе резерв роста эффективности использования двух других ресурсов, и, как следствие, уменьшение количества используемой материи и энергии.

Наряду с понятием «информация» распространение получило и понятие «данные». «Человечество столкнулось с кризисом, который само же создало: оно тонет в море информации. Мы произвели на свет столько статистических данных, формул, образов, документов и деклараций, что не в состоянии их усвоить. И вместо того, чтобы искать новые пути осмысления и усвоения уже созданного, мы все более быстрыми темпами продолжаем производить новую информацию. За последние десятилетия мы создали горы данных, которые, однако, не породили ни единой мысли. Вероятно, такого рода данные следовало бы называть не информацией[2], а эксформацией, поскольку они существуют сами по себе, вне человеческого мозга и независимо от него. … Большие объёмы неиспользуемой информации в конечном счёте превращаются в своего рода загрязнение окружающей среды. … Мы наблюдаем кризис образования на фоне избытка информации — и это не простое совпадение. Образование — это процесс переработки знаний, однако нам кажется проще генерировать новые факты, чем хранить и использовать те, которыми мы уже обладаем. Мы же, сталкиваясь с невежеством, производим ещё больше информации, не

[2] Термин «информация» возник путём слияния предлога «in» и корня «forme». Он означал нечто упорядочивающее, придающее форму. Информатором называли домашнего учителя, а информацией – учение, наставление. Леонов В.П. Современные проблемы информатики. Введение в семиотику информационных технологий: Учебное пособие. – Томск: Изд-во НТЛ, 2011. – 248 с.

желая понять, что хотя она, возможно, имеет ценность, но заменить знания, тем более мудрость ей не дано. И производя исходную информацию в гораздо больших количествах, чем когда-либо, мы начали вмешиваться в процесс превращения ее в знания» [15]. Ещё за 30 лет до этой статьи известный писатель и философ Станислав Лем, в своей прекрасной книге «Сумма технологии» [28] писал об этом явлении, называя его Мегабитовой бомбой.

Последний аспект, о котором в заключении говорит Альберт Гор [15], весьма актуален для тех структур, в которых наблюдается постоянный рост объёмов первичной информации. Например, расширение тематики научных исследований приводит к экспоненциальному росту объёмов собираемых данных. Что, в свою очередь, приводит к уменьшению глубины анализа этих данных. Например, как следствие этого в отечественных медицинских публикациях в последние годы наблюдается рост статистической вампукизации [42]. Отметим, что данное явление не специфично лишь для кардиологии. Данная проблема приобретает всё большие размеры во всех сферах деятельности. И как следствие этого появилось новое научное направление – информационная экология[3].

Одной из основных причин, провоцирующих такое снижение уровня использования, анализа первичных данных, является их доступность. В качестве примера приведём доступность в настоящее время в интернете колоссальных объёмов электронных публикаций, музыкальных записей, видеофильмов, программного обеспечения, и т.д. В результате посетители сайтов с этими информационными ресурсами выкачивают сотни, тысячи гигабайт информации. Однако при этом доля читаемых ими книг, просматриваемых фильмов, прослушиваемой музыки, и используемых программ, систематически снижается.

Итак, сегодня информация является одним из важнейших ресурсов развития науки и всего общества, наряду с материальными, энергетическими и людскими ресурсами. Однако реальную ценность имеет лишь структурированная информация, получаемая путём анализа исходных данных, превращаемая этим самым в ЗНАНИЕ. Например, проводимый врачом анализ данных по анамнезу конкретного пациента, в совокупности с результатами других анализов, позволяет ему сформулировать более точный диагноз и выбрать наиболее оптимальную тактику лечения конкретного пациента, нежели простое пролистывание историй болезни десятков пациентов.

В этой связи необходимо понимать, что «…понятия **информации и знания** следует отделять друг от друга. Знания являются переработанной информацией, они отражают связь между явлениями, выявленные закономерности и отвечают на вопросы «как?», «почему?», «что будет, если…?» и т. п., в то время как информация даёт ответы лишь на такие вопросы: «что?», «кто?», «когда?», «где?». Знание – это средство снижения неопределённости и риска, способствующее достижению определённых целей. Очень точно этот аспект отразил английский учёный, основоположник эмпиризма, Френсис Бэкон в своём афоризме «Знание – сила». Таким образом, знание есть продукт концентрации информации. Причём не просто концентрации в пространстве и времени, но ещё и концентрации в новом формате. В частности, в виде конкретных

[3] Мизинцева М.Ф., Королева Л.М., Бондарь В.В. Информационная экология. – М., 2000. – 231 с.

выражений отражающих исследуемые зависимости. Истинное Знание, по Бэкону, достигается с помощью познания причин, либо установления взаимосвязей между отдельными факторами, признаками, свойствами. Истинное Знание, в конечном счёте, необходимо для его носителей как средство увеличения возможностей выживания, как отдельных особей, так и всего сообщества в целом. Ошибочное же знание, есть не что иное, как дезинформация, уменьшающее вероятность выживания, как конкретной особи, так и сообщества в целом. Применительно к медицине это означает, что ошибочные медицинские технологии приводят к росту потребления неэффективных технологий, псевдолекарств, росту смертности вследствие роста доли ошибочных диагнозов, и т.д. И как итоговый результат всех этих действий к росту смертности населения, снижению рождаемости и вымиранию населения государства и его распаду» [46].

Поскольку знание является результатом переработки исходных данных, требующей определённых затрат, то очевидно, что здесь можно сравнивать СТОИМОСТЬ ДАННЫХ и СТОИМОСТЬ ЗНАНИЯ. Уточним, что речь идёт именно о СТОИМОСТИ получения этого знания, а не о его ЦЕННОСТИ (стоимость и ценность – разные понятия). Однако стоимость получения этого ЗНАНИЯ зависит как от технологии его извлечения из информации, так и уровня профессионализма специалиста, который извлекает это знание, и производительности используемой технологии извлечения знания. Например, если статистическим анализом занимается не профессионал, а дилетант, то очевидно, что производительность такого труда будет минимальна, а стоимость полученного дилетантом знания будет максимальна. Более того, нет никакой уверенности в том, что полученное им знание действительно будет истинным, надёжным, и, следовательно, ценным, т.е. очень полезным для его применения. Аналогичным образом, один и тот же метод статистического анализа может быть реализован путём проведения соответствующих вычислений, например, с помощью ручки или карандаша, и бумаги. А может быть реализован и с помощью таких пакетов, как EXCEL, STATISTICA, BioStat, GenStat, Mathcad, MedCalc, NCSS, Origin, SAS, JMP, SPSS, SYSTAT, Statgraphics, Stata, StatPlus, StatXact, Systat, MATLAB, MINITAB, ORIGIN, Cytel Studio, StatXact, LogXact, STADIA, R, S-plus, Epi Info, и т.д. Однако и в этом случае производительность и корректность полученного знания будет определяться опять же тем, кто будет проводить данный анализ: дилетант, или же профессиональный биостатистик. Иллюстрацией такой непроизводительной и дорогостоящей технологии получения ненадёжного знания, как раз и является выполнение статистического анализа биомедицинских данных не профессиональными биостатистиками, а медиками, не имеющими специальных знаний в этой области. При этом, чем более ценным и надёжным будет знание, полученное при анализе исходных данных, тем дольше оно будет использоваться в практике и цитироваться в последующих публикациях.

Не следует также игнорировать и следующий аспект получения исходных данных. Как правило, исходные данные есть результаты различных анализов (состав крови, ЭКГ, анализ мочи, мазки на флору, микробиологические посевы, ДНК-диагностика, биохимический и иммунологический анализ, цитологический анализ, и т.д.) производимых специалистами с использованием большого числа вспомогательного оборудования и расходных материалов. Т.е. в

совокупности СЕБЕСТОИМОСТЬ ИСХОДНЫХ ДАННЫХ достаточно высока. И поэтому из них должно извлекаться максимальное количество ценной информации.

Одним из признаков качества и ценности публикуемого нового знания является «факт уменьшения интенсивности использования документов по мере увеличения времени, прошедшего от момента публикации документа. Интенсивность же использования документов обычно выявляется путем установления частоты цитирования или справочно-библиографических запросов. ... В качестве показателя интенсивности старения информации широко используют такой показатель, как полупериод жизни научных статей. Полупериодом жизни статей называется время, в течение которого опубликована половина всей цитируемой в настоящее время литературы по конкретной области знания» [43]. Применительно к рассмотренным выше статьям можно сказать, что простейшие процедуры сравнения групповых средних, это ещё не «вершина познания». И в силу ненадёжности и сомнительности подобного знания их использование не приведёт к весьма существенным эффектам. Вследствие чего полупериод жизни таких публикаций будет существенно короче, чем у публикаций с надёжными результатами статистического анализа.

Отметим при этом ещё и следующий аспект. Уже давно установлено наличие отрицательной корреляции между частотой цитирования публикаций, и численностью соавторов публикации. То есть чем больше число соавторов статьи, тем реже её цитируют в других статьях. Такая зависимость вполне объяснима. Поскольку наличие многих соавторов автоматически приводит к дискретизации, несвязанности отдельных фрагментов публикации, так как эти фрагменты написаны разными соавторами. И такая структура публикации усложняет её восприятие читателями, тем самым вызывая снижение интереса к подобным публикациям. Не случайно в зарубежных публикациях, учитывая сей аспект, в конце статей можно прочитать выражение авторской благодарности тем своим коллегам, которые оказали им косвенное содействие в получении публикуемой информации, но при этом не были включены в соавторы. Что практически не встретишь в отечественных статьях. Другой же причиной увеличения численности соавторов часто является практикуемая во многих организациях практика дописывания в соавторы «свадебных генералов», т.е. своих руководителей, в надежде, что при таком авторском составе их публикация быстрее увидит свет.

Поскольку к имеющимся исходным данным можно многократно применять разнообразные, всё более сложные и эффективные методы анализа, то ЦЕННОСТЬ исходных данных во многом определяется как ЦЕННОСТЬЮ получаемых из них ЗНАНИЙ, так и кратностью анализа этих данных. Однако со временем снижается не только ценность публикаций, но и ценность любых исходных данных. В итоге снижается и целесообразность дальнейшего анализа этих данных. Учитывая интенсивное развитие разнообразных методов многократного анализа данных, за рубежом всё больше усилий вкладывается в создание огромных «хранилищ данных» (data warehouse), включая и так называемые «облачные хранилища данных».

В это связи уместно упомянуть о том, что согласно действующему законодательству исходные данные, оформленные в виде баз данных (БД), могут быть зарегистрированы, и по

усмотрению их владельца, доступ к ним могут иметь многие лица. Например, те данные, результаты анализа которых приведены в публикуемых статьях, через некоторое время следует делать доступными для многих иногородних специалистов, работающих в этой же области. Это могут быть не только профессиональные медики из других НИИ и медицинских вузов, но также и студенты этих вузов. Работая с подобными базами данных, они могут овладевать технологиями их анализа. В качестве примера таких БД, можно назвать известную электронную библиотеку данных DASL (http://lib.stat.cmu.edu/DASL/) и OzDASL - Australasian Data and Story Library (http://www.statsci.org/data/index.html) Именно по образцу этих БД мы создали RusDASL - российскую библиотеку данных для изучающих биостатистику. (http://www.biometrica.tomsk.ru/rus_dasl.htm) Статистика обращаемости к этим файлам, сгруппированным по методам статистического анализа, показывает постоянный рост количества обращений к ним.

Повышение уровня требований к качеству информационной продукции в экономически развитых странах наблюдается практически во всех отраслях, начиная от оборонных отраслей и заканчивая гуманитарными науками. В качестве типичной реакции на эту тенденцию, можно указать на фармацию. «Развитые страны предъявляют, как правило, более высокие требования в отношении качества производимой продукции. Так, введение чрезвычайно жестких норм на предварительную апробацию медицинских препаратов в США в 70-е годы вынудило многие фармацевтические корпорации вынести научные исследования за рубеж, в основном в развивающиеся страны. Не удивительно, что за 1974-1977 годы рост расходов на зарубежные НИОКР американских фармацевтических ТНК превысил 200% по сравнению с 34% внутри страны» [48]. А вот к чему приводит крайне низкий уровень подобных требований в отечественной фармации. 16 апреля 2013 г. мне пришло письмо от сотрудницы московской фармкомпании Полины Боркиной следующего содержания. «Знаю, что вы собираете интересные случаи описания статистики. Хочу с Вами поделиться своей находкой. Данное определение нашла в тексте статистического отчёта, предоставленного нашей компании по результатам проведённого для нас клинического исследования профиля безопасности лекарственного средства». Вот фрагмент этого отчёта:

> К параметрам с нормальным распределением были отнесены непрерывные данные, т.е. данные которые измеряются в абсолютных величинах. Дискретные показатели, не имеющие дробной части, относились к данным с непараметрическим распределением.

Абсурдность подобного утверждения очевидна. Во-первых, не все «непрерывные данные» (т.е. признаки) имеют нормальный закон распределения вероятностей. Во-вторых, не бывает «непараметрических распределений» вероятности. ВСЕ (NB!) законы распределения вероятностей имеют ПАРАМЕТРЫ. В т. ч. и те законы распределения, которые не являются нормальными. Исполнители же этого клинического исследования, выполняя требования производителя лекарств, такими абсурдными результатами будут аргументировать безопасность изучаемых лекарственных

средств. Очевидно, что подобных исполнителей здоровье будущих потребителей данных лекарств не волнует. На первом месте у них стоит прибыль от продажи подобных лекарств, а не здоровье потребителей этих лекарств. Поскольку ни Минздрав, ни другие организации, задействованные в этом фармацевтическом бизнесе, не предпринимают мер по недопущению подобного абсурда, то можно констатировать, что и их также не волнуют печальные последствия подобных нелепостей.

Понимание экономической эффективности продуктивного использования собираемых медицинских БД в настоящее время хорошо осознаётся и экономистами, работающими в медицинских НИИ. В качестве примера отличной публикации на эту тему, рассмотрим статью «Методы исследования операций и когнитивного анализа данных в решении задач лечебно-профилактических учреждений» [ПКК, № 4, 2011, с. 77-82]. Авторами статьи являются Дюбанов В.В. – ведущий инженер Института математики им. С.Л. Соболева СО РАН (Новосибирск), Руднев А.С. – к. ф-м. н., ведущий инженер Института математики им. С.Л. Соболева СО РАН (Новосибирск), Павловский Е.Н – к. ф-м. н., научный сотрудник Института дискретной математики и информатики Новосибирского государственного университета, Зозуля Ю.В. – доктор экономических наук, заместитель директора по экономике и развитию ФГБУ «ННИИПК им. акад. Е.Н. Мешалкина» Минздравсоцразвития России (Новосибирск), Самочернова А.С. – начальник финансово-аналитического отдела ФГБУ «ННИИПК им. акад. Е.Н. Мешалкина» Минздравсоцразвития России (Новосибирск), Сандер Д.С. – заместитель начальника финансово-аналитического отдела ФГБУ «ННИИПК им. акад. Е.Н. Мешалкина» Минздравсоцразвития России (Новосибирск). Вот что пишут 3 математика и 3 экономиста в своей статье. «Большинство стран мира сегодня предпринимает меры по переводу услуг здравоохранения на новый качественный уровень с целью сделать медицинскую помощь недорогой, общедоступной и более качественной за счет увеличения согласованности действий, улучшения диагностики, снижения издержек лечебно-профилактических учреждений. **Тщательный анализ поступающих потоков информации и эффективное управление ресурсами представляет собой наиболее перспективный путь развития в направлении решения озвученных проблем**. … В настоящее время не во всех медицинских учреждениях решена проблема автоматизации первичного учета данных. Однако даже в тех ЛПУ, где ведется учёт первичных данных в автоматизированном режиме, существующие решения … не закрывают всех актуальных потребностей в обработке поступающей информации (примеры таких задач будут рассмотрены далее). **Это происходит из-за того, что нарабатываемый большой объём данных требует мощных наукоемких аналитических решений** (в частности для верификации того, что первичный учет проводится корректно, и для многих других задач). Так, например, в ФГБУ «ННИИПК им. акад. Е.Н. Мешалкина» Минздравсоцразвития России в связи с осознанием этой ситуации было создано совместное предприятие по Федеральному закону от 02.08.2009 № 217-ФЗ «О внесении изменений в отдельные законодательные акты Российской Федерации по вопросам создания бюджетными научными и образовательными учреждениями хозяйственных обществ в целях практического применения (внедрения) результатов интеллектуальной деятельности» из группы исследователей (математиков, программистов, экономистов и медиков), в задачи которого входит создание нового продукта – аналитической (интеллектуальной) надстройки…» [47].

Далее авторы статьи приводят основные направления деятельности данной группы исследователей: использование теории решений для оптимизации технологий диагностики, создание эффективных расписаний, расчёт затрат на медикаменты и расходные материалы, оптимизация набора проводимых анализов, контроль качества оказания медицинских услуг, прогнозирование длительности этапов пребывания пациента в клинике, и т.д. «Последние десятилетия названы информационным бумом и характеризуются существенным ростом уровня автоматизации всех сфер приложения человеческих усилий. Многократно возросшие потоки разнородной информации – катастрофа для ручных методов обработки и учета, но благодатная почва для решения задач научного поиска, основанных на применении новейших достижений в области анализа данных. **Основной информацией, позволяющей решать широкий спектр задач анализа в здравоохранении, является статистика пролеченных пациентов**. … **Многочисленные мировые исследования показали, что в среднем от 15 до 20% проводимых диагностических тестов не являются необходимыми** [5]. Другими словами, издержки ЛПУ на этапе постановки диагноза могут быть существенно сокращены за счёт более разумного выбора схемы диагностики. Очевидно, ситуация принятия решения о том, какие именно тесты следует проводить на этапе, когда о пациенте известна только самая общая информация, характеризуется высокой степенью неопределённости. **Тем не менее первичной информации уже может быть достаточно, для того чтобы существенно сократить пространство выбора. Для этого следует обратиться к накопленной ранее статистике, выявить основные закономерности, например, между возрастом, полом пациента, регионом проживания, сезоном обращения, набором жалоб, результатами предварительных обследований и наиболее вероятным набором необходимых тестов, и применить их к новой ситуации**. Если результатов тестов оказалось недостаточно для постановки окончательного диагноза, то следующими стоит проводить те анализы, которые с большей вероятностью прояснят ситуацию. Если закономерности, выявленные в процессе анализа, достаточно устойчивы, ***что зависит от объёма статистики и от качества анализа***, то на множестве обследованных пациентов подобный подход будет оправдан. ***Выявление закономерностей, представленных в статистике, является одним из наиболее эффективных инструментов научного поиска.*** Традиционно, эмпирическое обнаружение законов – основа для дальнейшей работы по их аналитическому обоснованию. **Использование в медицине новых методов анализа для изучения накапливаемой статистики проведённых курсов лечения по множеству разнообразных факторов: диагноз, тактика лечения, состояние пациента, его данные (пол, возраст, наследственность), факторы окружающей среды и т. п. – может стать благодатной почвой для новых научных открытий**» [47].

По мнению авторов одним из важнейших направлений когнитивного анализа медицинских данных является выработка клинических рекомендаций. Вот что пишут об этом авторы статьи. «Анализ общей информации о пациенте может быть полезен не только на предмет предварительного заключения о том, какой диагноз наиболее вероятен, **но и для определения наиболее оптимальной стратегии лечения, например, с точки зрения минимизации рисков возможных осложнений**. Стоит понимать, что **цель анализа накопленной статистики по пролеченным пациентам** состоит не в том, чтобы вместо врача лечить пациента, подобный уровень автоматизации вряд ли допустим, **а в том, чтобы предоставить врачу простой и**

удобный доступ к опыту, накопленному специалистами многих лечебных учреждений. Это позволит ему сосредоточиться на наиболее значимых моментах, требующих его экспертного внимания, и не упустить из виду некоторых не типичных, но существенных в данном конкретном случае нюансов» [47]. В данном фрагменте сформулирована самая главная цель проведения анализа собираемой о пациентах информации: выбор оптимальной стратегии лечения, и обеспечение доступа врачам к полученным результатам анализа.

Повышение производительности труда за счёт использования современных информационных технологий, о чём пишут авторы данной статьи, в настоящее время становится в здравоохранении еще более актуальным ещё и по другой причине. В 2016 году Минфином планируется снижение федеральных расходов на здравоохранение. Так в 2014 г. планируется выделить 515 млрд руб., а в 2016 году уже только 360 млрд руб. Т.е. снижение составит 34%. А с учётом инфляции снижение будет ещё более значительным. Значит, к этому времени необходимо в полной мере реализовать предлагаемые авторами меры. Отметим, что задача выбора «**наиболее оптимальной стратегии лечения**», о которой выше сказали авторы статьи, достаточно продуктивно может решаться, в том числе, с помощью метода логистической регрессии [44]. Поскольку используя полученное уравнение логистической регрессии, можно, и изменяя в нужном направлении те или иные показатели, можно тем самым повышать вероятность благоприятного исхода для конкретного пациента.

Отсутствие изоляции как самих знаний, так и исходных данных, о чём выше пишут новосибирские математики и экономисты, является необходимым условием развития науки. «Любой ресурс, и в том числе и информация, лишь тогда даёт приращение иных ресурсов, когда он находится в использовании, обращении. А для такого обращения и необходим соответствующий транспорт. Афористично этот принцип можно выразить следующим образом: «Счастье не в обладании, счастье в использовании» [43]. Именно обращение Знания и способно приносить большие экономические эффекты. Экономика всего мира продуктивно использует знания, полученные учёными разных стран, применяя законы физики, химии, биологии, электротехники, радиолокации, генетики, и т.д. И чем шире и чаще использование этого знания, тем больше его экономический эффект. «Успехи в области науки и техники не возникают, да и не могут возникнуть, в изоляции. Они зависят от способности учёных и инженеров идти в ногу с последними достижениями, обмениваться информацией и учиться на опыте других. Как сказал когда-то Исаак Ньютон, «если я и видел дальше других, то только потому, что стоял на плечах гигантов». Не имея доступа к работам Николая Коперника, Галилео Галилей никогда не мог бы узнать, как использовать свои наблюдениям за спутниками Юпитера. Не зная об уравнениях поля, выведенных блестящим шотландским физиком XIX века Джеймсом Клерком Максвеллом, итальянский инженер двадцатого столетия Гульельмо Маркони не смог бы изобрести радио, а Альберт Эйнштейн не создал бы свою теорию относительности. Короче говоря, темпы технологического прогресса зависят от того, насколько широко доступны для членов общества знания и насколько люди могут делиться ими, **т.е. от уровня развития технологии обработки информации.** Таким образом, … **скорость, с которой развивается технология в**

обществе, определяется относительным уровнем его способности усваивать и обрабатывать информацию» [52].

Любой ресурс, включая информационные ресурсы в виде медицинских БД, представляет собой авансированные затраты. Рентабельность таких затрат есть отношение чистой прибыли к себестоимости затрат на их получение. В отличие от энергии и материальных ресурсов, которые при их использовании либо полностью, либо частично трансформируются и исчезают, информационный ресурс при его анализе не уменьшается. Действительно, например, при проведении нескольких видов статистического анализа один и тот же массив остаётся в прежнем своём объёме. Т.е. для информационного массива возможна (и необходима!) его многократная эксплуатация, при которой от первоначальных примитивных методов, типа сравнения средних, можно постоянно переходить к более сложным методам анализа. В этом случае максимум приращения количества нового знания в результате повторных аналитических операций достигается на некотором ограниченном количестве шагов. После которого новые анализы того же самого массива данных приносят всё меньшее количество знания. Однако масштабы такого приращения зависят от того, насколько профессионально выполняется статистический анализ этих исходных данных. Именно по этой причине медицинские научные организации должны иметь в своей структуре лаборатории биостатистики, а не церковь, как это сделал директор Томского НИИ кардиологии РАМН.

Современные экономические теории включают в производственную функцию, отражающую взаимосвязь объёма производимого продукта W с объёмами необходимого труда К и капитала L, также и объём используемой при этом информации I. Вот как выглядит наиболее популярная подобная производственная функция: $W(t) = aK^{\alpha}L^{\beta}e^{\gamma I}$, [21] где все константы положительны и α+β=1. «Такой подход автоматически предполагает по отношению к информации закон возрастающей отдачи, так как для данной функции независимо от значений входящих в неё параметров $\frac{\partial^2 W}{\partial I^2} > 0$.[4] Ряд исследователей в самом деле полагает, что применение информации в производственных процессах подчиняется закону возрастающей производительности. В качестве примера сошлемся на позицию Дэниела Белла: «Замена рабочих машинами приводит к экономии не только труда, но и инвестиций, так как каждая следующая единица капитала более эффективна и производительна, чем предыдущая, и, следовательно, на единицу продукции требуется меньше затрат…» [6].

Обсуждаемая новосибирскими математиками и экономистами НИИ им. акад. Е.Н. Мешалкина задача когнитивного анализа, в решении проблем медицинской науки и практической медицины, результативно решается во многих странах зарубежья. Так около 60 стран в мире имеют организации, проводящие оценку технологий здравоохранении (ОТЗ). При этом в штате этих организаций состоят врачи различных клинических специальностей, эпидемиологи,

[4] Выражение $\frac{\partial^2 W}{\partial I^2} > 0$ является частной производной второго порядка, свидетельствующей о том, что при постоянных значениях труда К и капитала L прирост информации I будет ВСЕГДА приводить к приросту объёма продукции W.

физиологи, экономисты, статистики, социологи [51]. «К сожалению, страны, в наибольшей мере нуждающиеся в систематической оценке медицинских технологий, в наименьшей степени в ней задействованы. Программ ОТЗ практически нет в африканских странах, во многих странах Юго-Восточной Азии и Латинской Америки. Отсутствие в России государственной организации, выполняющей ОТЗ, при наличии определённого опыта в области продвижения доказательной медицины, отмечается зарубежными экспертами в обзорных статьях с удивлением, и трактуется как факт, означающий недостаточную политическую поддержку этого направления развития здравоохранения» [1].

В отличие от государственных структур, необходимость оценки эффективности внедрения достижений медицинской науки в практическое здравоохранение, осознаётся российскими медиками. «На практике, когда в здравоохранении широко используется медицина, основанная на доказательствах, подчас сложно оценить экономическую эффективность внедрения конкретных результатов современных технологий диагностики, лечения, реабилитации и профилактики **без применения достоверных математических моделей**» [20]. Авторы данной статьи достаточно подробно описывают используемую ими технологию оценки эффективности, используя для этой цели, как обширный набор показателей, так и обширный набор методов анализа. Именно такое сочетание и даёт оценку глубины переработки имеющихся информационных ресурсов. «Для отечественных исследователей в области биомедицины это затруднено не только отсутствием в структуре исследовательских организаций РАМН и вузов лабораторий биостатистики, но и недостатком специализированной периодики, содержащей материалы для повышения уровня подготовки медиков и биологов в области применения методов статистики. Отметим, что в Англии ещё в 1901 г. появился такой известный журнал, как "Biometrika", в США в 1945г. журнал "Biometrics", в 1959 г. в Германии появился журнал "Biometrische Zeitschrift", затем появились "Психометрика", "Технометрика", "Эконометрика" и "Наукометрия", материалы которых посвящены применению статистики в различных областях науки. Все эти журналы не только выполняют обучающую функцию, но и прививают читателям вкус и потребность к грамотной статистической обработке экспериментальных данных. С 1982 г. издательством JOHN WILEY & SONS издаётся специализированный журнал "Statistics in Medicine" (http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1097-0258), а издательством ELSEVIER выпущено уже более 50 томов журнала "Computer Methods and Programs in Biomedicine". Несмотря на то, что подписка на эти журналы стоит порядка 1000 долларов США, библиотеки крупнейших западных университетов и медицинских центров выписывают их. Тогда как в России не только не выписывают их, но и практически не знают об их существовании. … Значительным тормозом в приобщении нашей страны к общемировой биостатистической культуре является и то, что Россия не является членом специализированных организаций созданных с целью расширения практики использования достижений статистики в биологии и медицине. Так, в 1938 г. вначале была создана Биометрическая секция американской статистической ассоциации. В 1947 г. в Вудс-Холе (США) ею была созвана Первая международная биометрическая конференция, организовавшая Международное биометрическое общество. Долгие годы единственным членом этого общества от СССР был А.А. Любищев. Много лет функционирует Международное общество клинической биостатистики, национальные

отделения которого есть не только в таких странах как Венгрия, Польша и Румыния, но и во многих африканских странах, не говоря уже о США и большинстве европейских стран. Однако Россия, не участвует в работе этих организаций» [29]. Отметим, что 24-28 августа 2014 г. в Вене (Австрия) состоится юбилейная 35 Конференция Международного общества клинической биостатистики (ISCB), которое было организовано в 1978 году. [http://me-kono.eu/conferences/35th-annual-conference-of-international-society-for-clinical-biostatistics-iscb-2014]

Глубина и качество переработки любого ресурса, будь то информационный ресурс, или же минеральное сырьё, как раз и определяют эффективность и стабильность любой экономики. Например, первоначально нефть использовалась как смазочный материал для трущихся механических устройств. Аналогично, изучая собранные кардиологические данные, можно ограничиваться лишь сравнением средних в группах. Однако по мере вложения энергии, информации и материи в исследование нефти, полученная новая информация о составе нефти позволила использовать её как сырьё для получения керосина и бензина для двигателей внутреннего сгорания. Что позволило существенно увеличить эффективность использования данного материального ресурса. Точно также используя на профессиональном уровне современные технологии биостатистики, из собранных кардиологических данных можно извлечь гораздо больший объём Знания, обращение которого в науке и практике даст соответствующий экономический эффект.

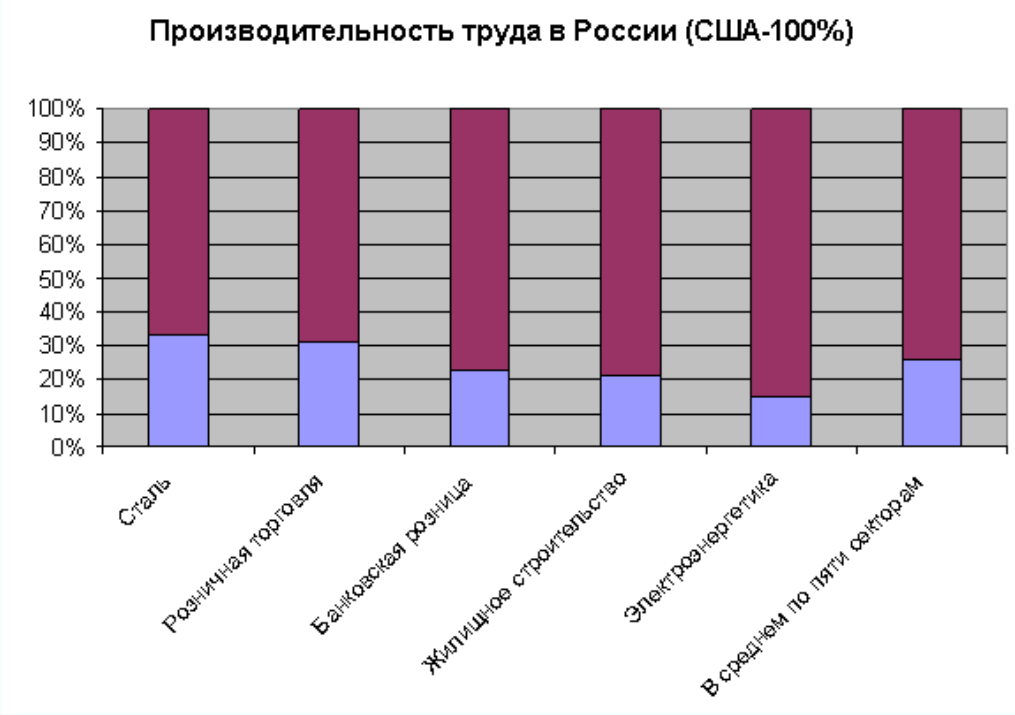
Аналогичная ситуация и в отечественной газовой промышленности. Однако, несмотря на то, что фактически вся экономика России держится на «нефтегазовой игле», уровень развития этой промышленности катастрофически отстаёт от мирового уровня. «Технический уровень большинства НПЗ также не соответствует передовому мировому уровню. В российской нефтепереработке основными проблемами отрасли, после низкого качества получаемых нефтепродуктов, остаются **низкая глубина переработки нефти** – (в России – 72%, в Европе – 85%, в США – 96%), отсталая структура производства – минимум вторичных процессов, и недостаточный уровень процессов, улучшающих качество получаемых продуктов. Ещё одна проблема – высокая степень износа основных фондов, и, как следствие, повышенный уровень энергопотребления. На российских НПЗ около половины всех печных агрегатов имеют КПД 50 – 60% при среднем показателе на зарубежных заводах – 90%» [64].

«Ранее страна потребляла 75% добываемой нефти и экспортировала 25%. Сейчас наоборот – внутри страны используется порядка 25% объема добываемой нефти и продуктов её переработки. Отечественная нефтепереработка в настоящее время характеризуется низкой рентабельностью, обусловленной недооснащённостью НПЗ современными процессами глубокой переработки нефти, и высокой изношенностью основных фондов. В стране уже в течение нескольких десятилетий не строятся новые НПЗ. …В 1990 г. уровень потребления нефтепродуктов на душу населения составлял 1,6 тонны, в настоящее время 800 кг. Наибольшую выгоду (в разы) получает та страна, которая потребляет нефть, а не торгует нефтью. Потребление продукции нефтепереработки и нефтехимии на душу населения – важнейший показатель, влияющий на рост ВВП и уровень жизни населения. Увеличение глубины переработки нефти одновременно решает задачу сырьевой базы для таких важных отраслей народного хозяйства, как нефтехимия и химия.

Прочитай и перешли своим коллегам…

Стоимость нефтехимической продукции в 5–10 раз дороже нефтепродуктов» [55]. Отметим, что при этом в отечественной нефтяной промышленности производительность труда ниже мирового уровня в 3-5 раз. Низкий уровень производительности труда в целом по России характерен не только для нефтегазовой промышленности.

Отставание России по производительности труда от Евросоюза и США в ряде отраслей достигает 30 раз, (данные Минэкономразвития, 2008 г)



Так отечественная ракетно-космическая промышленность ежегодно производит продукции из расчета 14,8 тысячи долларов на одного работающего, в ЕС этот показатель - 126,8 тысячи долларов, в США - 493,5 тысячи долларов (т.е. выше в 33,3 раза). В судостроении России на производство 1 тонны продукции тратят в три раза больше рабочего времени, чем в Южной Корее. Автомобилей же на одного рабочего в российском автопроме выпускается в четыре раза меньше, чем в ЕС.

В целом в 2008 году отмечалось, что по производству добавленной стоимости на одного занятого Россия отстаёт от США в 5,9 раз, от стран G7 — в 5,2 раза, от Японии — в 4,3 раза. А по производству за один отработанный час разрыв ещё выше: с США — 6,1 раз, со странами G7 — 5,4 раза. В 2000 году производительность труда в России составляла 14,9 % от уровня США.

«В нефтегазовой сфере сотрудник российских предприятий в год производит продукции на 0,52 миллиона долларов, в то время как в американских компаниях сотрудник приносит до пяти миллионов долларов в год. Таким образом, эффективность труда российских сотрудников составляет лишь 10,4 процента от аналогичного показателя у американцев. В металлургии американец в год производит продукции на 570 тысяч долларов, а его российский коллега — лишь на 150 тысяч долларов. В электроэнергетике сотрудник российских компаний приносит предприятию 200 тысяч долларов, а в США — 790 тысяч долларов». Самыми неэффективными

компаниями в России назвали нефтегазовые и металлургические [56]. Одной из причин этого явления исследователи назвали неоптимальную организацию производства, т.е. организационную патологию. О необходимости повышения производительности труда в России говорил и её президент В. Путин, выступая 2 октября 2013 г. на инвестиционном форуме «Россия зовёт!».

Являясь составной частью экономики, производительность труда в медицине в таких условиях также не может иметь высокий уровень. Так агентство «Финмаркет», оценившее производительность в семи отраслях экономики России на основании данных Росстата и Всемирного банка показало, что самая минимальная производительность труда наблюдается в здравоохранении, образовании и социальной сфере. «В этом сегменте на каждого занятого производится услуг в среднем на 12,9 тысячи долларов в год. Это в шесть раз меньше, чем в США (76 тысяч долларов)»[62]. Одним из проявлений этого является следующий факт. В ноябре 2006 г. в Москве проходил I национальный конгресс терапевтов «Новый курс: консолидация усилий по охране здоровья нации». Выступая на нём, академик РАМН А.Г. Чучалин сказал, что практически каждый третий диагноз ставится российскими врачами неверно, тогда как в США врачебные ошибки составляют 3–4 %, в Великобритании – 5 %, во Франции – 3 %.

Какие же факторы влияют на производительность труда? В первую очередь это технологии. Бытовое представление о технологии сводится в основном к машинам и механизмам, т.е. орудиям труда. **Но для их создания, как и для работы с ними, необходимы специальные знания, а также наличие соответствующей инфраструктуры. Т.е. технология – это информация, устройства и системы, повышающие производительность труда.** Высокая производительность труда обеспечивается соответствующим уровнем технологий, которые, в свою очередь, требуют высоких уровней знаний и квалификаций. Поэтому первоочередными задачами повышения производительности труда в медицине является повышение уровня специальных знаний и навыков работы с современными технологиями, в т. ч. со статистическими технологиями. Фактически на этом и базируется концепция ДМ. «Производительность труда, это, … самое важное, самое главное для победы нового общественного строя. В.И. Ленин. Великий почин. ПСС, т.39, с.22». Продолжая эту мысль, можно констатировать, что снижение производительности труда в различных отраслях России, в т. ч. в медицине, это путь к ликвидации существующего государственного устройства, с целью повышения производительности труда.

Появление новых технологий и методов влечёт за собой обязательное изменение организационной структуры. При этом для функционирования этих структур необходимо привлечение соответствующих специалистов. Например, довольно длительное время в медицинских вузах и НИИ, несмотря на увеличение количества компьютеров, отсутствовали специальные службы отладки, ремонта и настройки этой вычислительной техники. Однако появление этих служб через некоторое время связано также и с тем, что в период обучения врачей в вузах им не читаются специальные курсы об архитектуре компьютеров, о системах программирования, ремонте этой техники, и т.д. В итоге, для обслуживания компьютеров, томографов, и другого сложного оборудования, в штатное расписание вузов и НИИ вводят соответствующих специалистов. Но поскольку будущим медикам в вузах читают примитивные курсы по статистике, то после окончания вузов у них сохраняется представление о том, что

статистика – это просто, легко, и доступно. И поэтому выполнять статистический анализ можно самостоятельно. Другой причиной отторжения профессиональных биостатистиков от выполнения статистического анализа собираемых данных, является и нежелание расширять списки соавторов публикаций путём включения в них биостатистиков. В итоге возникает то, что современной терминологии называется «информационный луддизм». Т.е. отвержение сложных аналитических технологий вместе со всеми сопутствующими компонентами.

Анализ проблемы снижения производительности труда, проведённый ведущими экономистами страны, установил, что доминирующей причиной столь низкой производительности труда в России является неэффективная организация труда. В экономике появился даже такой термин, как «**организационная патология**», введённый в 1972 году польской исследовательницей Ядвигой Станишкис в рамках анализа организационных структур и их дисфункций. [http://ru.wikipedia.org/wiki/__Организационная_патология] [2] Это понятие, близкое к медицинскому термину, трактуется как отклонение от нормы и как дисфункция. Проявление отклонений от нормы может иметь самые разные виды. Например, в статье «Вопрос не только об автомобилях, но и об унитазах» на LENTA.RU от 11 сентября 2013 г.(http://lenta.ru/articles/2013/09/11/roi/), читаем: «Я на совещании в Генпрокуратуре побывала в туалете — там **подогреваемые японские унитазы с пультом управления и инструкцией из 16 пунктов, как этим унитазом пользоваться. Он 15 тысяч долларов стоит.** Зачем такой унитаз Генпрокуратуре, непонятно. Поэтому вопрос не только об автомобилях, но и об унитазах тоже», — заключила Яковлева, вызвав бурную реакцию собравшихся». Другой пример. «В сфере госзакупок доходит до анекдотичных примеров - ёршики для унитазов по 12 тыс. руб. за штуку (мэрия Санкт-Петербурга), комплекты мебели за 4,4 млн (МВД) и ремонт туалета за 1,6 млн (мэрия Красноярска)» [59]. Данные примеры наглядно демонстрируют, как организационная патология приводит к неэффективному использованию финансовых ресурсов. Столь же неэффективно можно использовать и информационные ресурсы в кардиологии, ограничивая проводимые исследования лишь сравнением средних, игнорируя при этом применение современных методов биостатистики. Фактически и в этом случае мы получаем слишком высокую себестоимость ответа на вопрос: отличаются ли статистически значимо средние величины в группах сравнения.

Достаточно подробно проблему организационной патологии описал в своём интервью «Россия на игле» (http://www.mk.ru/politics/interview/2013/11/14/945585-rossiya-na-igle.html) бывший вице-премьер и министр финансов Алексей Кудрин. На вопрос корреспондента каков рецепт улучшения экономической ситуации в России, он ответил: «– Если его описать самыми общими словами, то, простите за банальность, это рост производительности труда. А это достигается целой цепочкой мер. Экономике России жизненно необходимо массированное внедрение новых технологий».

Организационная патология, применительно к медицинской науке, проявляется не только в отсутствии в медицинских НИИ и вузах служб статистического сервиса в виде лабораторий биостатистики. Более простой формой такой организационной патологии является и отсутствие отраслевых стандартов на редакционные требования медицинских журналов, о чём мы упоминали

выше. Разумеется, реализация таких стандартов возможна лишь при условии понимания важности данных аспектов в развитии медицинской науки руководителями данных соответствующих, и при активном участим в их разработке организаций, проповедующих доказательную медицину. За рубежом уже давно осознали важность этого аспекта медицинской науки. И на основе этого понимания в 1978 г. был создан Международный Комитет Редакторов Медицинских Журналов (International Committee of Medical Journal Editors, ICMJE, http://www.icmje.org/).

«К чему приводит отсутствие подобных стандартов и лабораторий биостатистики можно проиллюстрировать на многочисленных примерах клинических испытаний биологически активных добавок (БАД) к пище. Несколько лет назад один из моих корреспондентов прислал мне книгу «Результаты клинических испытаний БАД к пище» написанную Департаментом исследования качества жизни и питания компании Vision International People Group. Книга содержит результаты 14 клинических испытаний различных БАД, выполненные следующими организациями: 1) НИИ питания РАМН, 2) Российской медицинской академией последипломного образования Министерства здравоохранения и социального развития РФ, 3) Санкт-Петербургской государственной академией физической культуры имени П.Ф. Лесгафта, 4) Московским областным научно-исследовательским институтом акушерства и гинекологии МОНИИАГ, 5) Российским онкологическим научным центром РАМН имени Н.Н. Блохина, и 6) Государственным научно-исследовательским институтом стандартизации и контроля медицинских биологических препаратов имени А.А. Тарасевича. Отчёты об итогах клинических испытаний БАД в этих 6 организациях утверждены их руководителями и заверены печатями. Интервал дат утверждения этих отчётов от 1998 г. до 2005 г. … Однако, несмотря на различие организаций, в которых проводились эти испытания, различие руководителей и дат, эти отчёты имеют формулировки, дословно совпадающие в описании статистических аспектов этих испытаний. Нетрудно догадаться, что означает этот факт» [45].

Завершая данный раздел, можно констатировать, что помимо наличия внутренних проблем медицинских НИИ и вузов по гармонизации накопления исходных данных и анализа, о чём писал в своей статье Д.С. Саркисов [57], присутствует и системная составляющая, препятствующая решению этой проблемы.

Заключение

Где необходимость,
там и возможность

_____
Пифагор

Наукометрический анализ медицинских публикаций второй половины 20-го века показал, что в отечественной медицинской науке доминирующей является сдвиговая парадигма

[39]. Это проявляется в преобладании в публикациях статистических процедур сравнения групповых средних. Наглядно это можно представить как проверку гипотез о том, что температура тела популяции больных выше температуры тела популяции здоровых. Возрастающие возможности современных методов биостатистики приводят к росту потенциальных возможностей их использования в медицинской науке. Так автор этих строк, занимаясь биостатистикой уже более 30 лет, все эти годы осваивал новые методы статистического анализа. И по сей день появляются новые методы биостатистики, позволяющие извлекать из биомедицинских данных весьма ценное знание. Системный подход в медицинской науке, реализующий упомянутую Д.С. Саркисовым гармоничность накопления фактов и их обобщения, возможен только при переходе от сдвиговой парадигмы к парадигме «человек = система». Это означает, что кроме вычисления средних величин в группах и их сравнения, необходимо учитывать и такую важнейшую компоненту, как **структуру взаимных связей между всеми изучаемыми признаками**. А для этой цели необходимо использование разнообразных многомерных методов биостатистики.

В 50-70-е годы прошлого века активную деятельность по продвижению статистики в биологию и медицину вёл Николай Александрович Плохинский, который в 1928 году начал свою работу в МГУ. С 1958 по 1962 год работал в Институте цитологии и генетики (Новосибирск), и затем вновь вернулся в МГУ. Первую свою книгу о применении статистики в биологии он издал ещё в 1937 году. Затем им были написаны ещё несколько книг на эту же тему. Наиболее известны из них такие книги, как «Алгоритмы биометрии», «Математические методы в биологии», и «Биометрия», на которых училось применять статистику в биологии и медицине не одно поколение учёных. Вот как начинается раздел «Введение» в книге «Биометрия» (2-е издание. Издательство Московского университета. 1970. – 367 с.).

«Современная биология не может развиваться без математической помощи. … Математика необходима для исчерпывающего извлечения информации о типичных объектах, разнообразии их, структуре этого разнообразия, о системах биологических взаимоотношений и взаимодействий, о разных биоценозах, о влиянии разных факторов на биологические объекты, развивающиеся в различных условиях. Чем обширнее и глубже вскрываются эти явления, тем большую силу получает человек при использовании природных богатств и, главное, при направленных преобразованиях живой природы. … Имеются такие задачи современной биологии, которые не только не могут быть решены без математики, но и трудно понимаемы без соответствующей математической подготовки. … Современный биолог должен знать основы математики. Многолетний обширный опыт использования математики в биологии выявил, кроме того, и формы наиболее успешного сотрудничества между биологами и математиками и наиболее эффективный метод внедрения математики в биологию. … Ещё на заре советской биометрии в 1925 г. биолог академик А.С. Серебровский определял отношения биологии и математики следующим образом… «**Биолог должен оставаться прежде всего биологом и отнюдь не превращаться в математика**. История статистических исследований в биологии показала, что только там, где исследуемый материал наряду с анализом математическим подвергался интенсивному биологическому анализу – только там метод дал положительные результаты,

продвинув биологов за границы доступного «невооружённому глазу». Безусловно, статистический метод получает права гражданства в биологических дисциплинах. Это внедрение в новые для него области предъявляет самому математическому методы новые требования, удовлетворить которые должны сами математики. И можно не сомневаться в том, что подобное сотрудничество биологов с математиками принесёт ещё богатые плоды. … Биология и математика образуют не механическую смесь методов, а качественно новый сплав, в котором оба элемента начинают выступать единым мощным фронтом. На определённом этапе развития этого синтеза комплекс новых методов приобретает свои специфические отличи и выделяется в особую отрасль знания. На стыке биологии и математики зарождается новая синтетическая наука с собственным предметом изучения, с особыми основными методами, особыми терминологией, символикой и специфическими задачами. Если отбросить **очень неудачный термин «вариационная статистика»**, то лучше использовать для обозначения этой новой науки более подходящее название – «биометрия», принятое в настоящее время большинством биологов».

К сожалению, даже в тех публикациях, где в списке литературы упоминается эта книга Н.А. Плохинского, можно прочитать и о том, что авторы использовали «метод вариационной статистики».

Необходимость активного продвижения статистики в биологию и медицину в то же самое время прекрасно понимали многие советские учёные. Например, академик А.Н. Колмогоров в своё время даже предлагал Министерству высшего образования СССР ввести в университетах подготовку студентов по специальности биостатистика. Однако чиновники министерства отказали ему в этом. Учитывая уровень деятельности и нынешних государственных органов управления, вряд ли можно рассчитывать, что в ближайшие годы будут приняты реальные мер по повышению качества отечественной науки. Подтверждением этого служат наблюдаемые в настоящее время примитивные реформы Академии наук, ВАК и технологии присуждения учёных степеней. Между тем актуальность этих мер с каждым годом повышается. О необходимости реализации этих мер в медицинской науке говорится, в частности, в Ереванской Декларации [23], принятой на международной конференции по доказательной медицине, состоявшейся 18-20 октября 2012 года в г. Ереване.

История развития науки и университетского образования показывает, что именно концентрация интеллектуальной мощи приводила к созданию первых университетов. Причём большинство университетов возникало благодаря самоорганизации сообществ интеллектуалов. И лишь потом их статус и структура становились объектами указов и положений внешних органов управления. Поэтому ликвидация организационной патологии в части игнорирования современных методов биостатистики возможна лишь при достижении руководящим звеном медицинских НИИ и вузов определённого уровня знаний в биостатистике. Оптимальным средством для этого является проведение специализированных семинаров по биостатистике для руководящего звена. Наряду с этим необходимо создавать в этих организациях и систему «маяков» – отличных аналитических исследований, результаты которых могут быть получены с помощью современных методов биостатистики. Образцы таких работ, оформленные в виде серии

статей, диссертаций, монографий, и доступные многим, будут служить для остальных исследователей этих организаций ориентирами, примерами для подражания.

В феврале 2013 г. для создания системы «маяков» (примеров для подражания с хорошим использованием статистики в биомедицине) мы объявили конкурс на бесплатное выполнение статистического анализа для медицинских НИИ и вузов. Согласно условиям конкурса заявки должны были подаваться от зам. директоров по науке НИИ и проректоров на науке вузов. Однако поступила всего лишь несколько заявок от руководителей таких организаций. Однако ни одна из них не отвечала поставленным условиям. При этом было очень много заявок от самих аспирантов и докторантов. Из переписки с этими аспирантами и докторантами, обучающимися или работающими в таких организациях, стало ясно, почему эта идея не заинтересовала их руководителей. Наиболее вероятная причина, которую сформулировал один моих корреспондентов, заключается в том, что наличие подобных примеров для подражания может снизить количество диссертаций подаваемых к защите в диссертационные советы этих вузов и НИИ. Что, вероятно, не очень выгодно для их руководителей.

Факт отставания России по уровню технологий от экономически развитых стран признаётся руководством страны. В решении подобных проблем имеют опыт многие страны. Приведём небольшой фрагмент из статьи "Россия по уровню технологий отстает от развитых стран на 50 лет".

"Сейчас только ленивый не пинает действующее правительство за бездеятельность и непрофессионализм. Наверное, чуть ли не каждый министр или вице-премьер в нём – сам по себе человек умный, толковый, симпатичный. Но вместе они не производят впечатление слаженной, ответственной команды. В том числе из-за отсутствия у правительства чётких, понятных всем целей (ориентированных, в первую очередь, на повышение благополучия граждан), отсутствия ответственности за достижение запланированных результатов. Нет ответственности, значит, нет никакого контроля. В отсутствии контроля – нет результата. Прописные, вроде бы, истины… Но нам-то от этого не легче.

Уже сейчас по уровню развития технологий мы отстаем от развитых стран примерно на 50 лет. Это ни для кого не секрет. Как, слава Богу, и то, что - ни мы первые, ни мы последние. И до нас были страны с такой дикой технологической задержкой в развитии. Поэтому накопился опыт решения такой проблемы. Взять хотя бы Финляндию. До революции 1917 года была частью российской империи, отсталым регионом. Став самостоятельным государством, они сделали ставку на развитие лесной промышленности. И вроде бы все получалось (качественная продукция, рост экспорта), но встать вровень с развитыми странами – ну, никак! Тогда чиновники из финского правительства собрали ученых, специалистов, видных политиков, предпринимателей. Почесали головы (какие же отрасли им необходимо развивать, чтобы не отстать от Западной Европы?) И тогда, исходя из масштабов страны, потребностей своей экономики, желаемого уровня благосостояния граждан решили сконцентрировать все свои усилия на телекоммуникационной отрасли. Вскоре появились ныне широкоизвестные мобильные телефоны Nokia. Решение, повторюсь, было принято загодя – примерно в 1980-м… В то время компания

Nokia была малоизвестным производителем резино-технической продукции - от галош до шин. Сейчас это одна крупнейших корпораций в мире. В начале нулевых финны снова устроили мозговую атаку и приняли решение сосредоточиться на развитии биотехнологий. Не сомневаюсь, что по прошествии 10 лет Финляндия и здесь станет ключевым игроком. А где будем мы, наша страна?"

Очевидно, что решение проанализированных выше проблем может иметь разные сроки реализации. Например, достаточно быстро могут быть разработаны новые редакционные требования журнала ПКК к присылаемым рукописям. Оперативность этого шага в настоящее время весьма актуальна, поскольку ВАК в скором времени направит в Минобрнауки свои предложения по формированию нового Перечня рецензируемых научных журналов и изданий, в которых будут публиковаться основные научные результаты диссертаций на соискание учёной степени. В результате чего количество изданий в этом Перечне сократится в несколько раз.

Учитывая тот факт, что количество электронных изданий неуклонно возрастает, целесообразно также и изменение формата медицинских журналов. В частности, можно наряду с изданием бумажных версий журналов, в которых имеются вполне объяснимые ограничения объёма публикаций, можно параллельно издавать и их электронные версии, в которых допустимые объёмы статей могут быть увеличены до 10 и более раз. Отмечу, что в настоящее время имеющиеся электронные версии журналов содержат статьи, идентичные по содержанию статьям в бумажных версиях. Недостатком этих электронных версий можно считать также тот факт, что чаще всего для прочтения отдельной статьи, читателю необходимо копировать весь конкретный выпуск этого журнала. Учитывая факт наличия не очень высоких скоростей копирования у российских пользователей интернетом, целесообразно сделать доступными для скачивания также и отдельные статьи этих выпусков. Расширенная электронная версия журналов позволит авторам статей очень подробно описывать достаточно сложные исследования, приводя при этом и весьма подробное изложение статистических аспектов.

Одним из признаков понимания руководством ПКК важности статистических аспектов публикуемых статей, по нашему мнению, должно быть наличие в новой версии редакционных требований достаточно конкретных и ясных формулировок по описанию статистики. Подобно тому, как это сделано в редакционных требованиях журнала «Экология человека». Следует также ввести в редакционные требования ПКК обязательное статистическое рецензирование рукописей независимыми специалистами в области биостатистики, а также включение таких специалистов в состав экспертного совета, принимающего решение о публикации статьи. Именно такая рекомендация и была приведена в "Ереванской Декларации". Полезно также и введение требования о предоставлении авторами статьи используемых в анализе данных, в том случае, если результаты статистического анализа будут оценены специалистами как сомнительные. Напомню, что такое требование имеется и в зарубежных журналах.

Несмотря на критический характер подобных обзоров, они всегда приводят к появлению положительного эффекта, как у их читателей, так и у авторов анализируемых публикаций. Например, если читатели данного обзора внимательно ознакомятся с авторскими составами

публикаций в разделе «Что такое ХОРОШО», и в разделе «И что такое ПЛОХО», то они увидят явную асимметрию этих составов, которая объясняется опубликованным нами несколько лет назад аналогичным региональным обзором.

P.S.

23 октября 2013 г. руководители НИИ им. Е.Н. Мешалкина получили 5 бандеролей с бумажными версиями данного обзора. В одном из сопроводительных писем к обзору, автор обратился к главному редактору журнала ПКК, д-р мед. наук, проф., академику РАМН А. М. Караськову, с предложением о публикации в ПКК краткой версии данного обзора. В случае согласия с этим предложением краткая версия данного обзора будет опубликована в журнале «Патология кровообращения и кардиохирургия». Однако ответа на это предложение пока не поступило... Учитывая актуальность обсуждаемой проблемы, бумажные версии обзора будут отправлены также руководителям Минздрава, Минобрнауки, ВАК, и ряду других руководителей, от которых зависит изменения сложившейся ситуации в лучшую сторону.

P.P.S.

Если удастся, то через 5 лет, в 2018 г., мы вновь проведём подобный обзор, что установить, насколько изменилась за прошедшие 5 лет ситуация с использованием статистики в кардиологии.

Литература

1. Авксентьева М.В., Омельяновский В.В. Международный опыт оценки технологий в здравоохранении. НИИ клинико-экономической экспертизы и фармэкономики РГМУ, Москва. Медицинские технологии. Оценка и выбор. 2010. №1, с. 52-58.
2. Ануфриева Н. И. Патология организации. Всероссийский экономический журнал ЭКО №12, 2006 г. http://econom.nsc.ru/eco/arhiv/ReadStatiy/2006_12/Anufrieva.htm
3. Амосов Н.М. Мысли и сердце. М.: Молодая гвардия, 1969. – 352 с.
4. Барт А. Г., Е. В. Вербицкая, В. Н. Солнцев. О состоянии дел и перспективах обучения статистическому анализу медицинских данных / Международный журнал медицинской практики. – 2006. – № 2. – С. 39–42. http://www.biometrica.tomsk.ru/bart.htm
5. Бащинский С. Е. Статистика умеет много гитик / Международный журнал медицинской практики. – 1998. – № 4. – С. 13–15. http://www.biometrica.tomsk.ru/gitiki.htm
6. Белл Д. Грядущее постиндустриальное общество: Опыт социального прогнозирования. М.: Academia, 1999. – 785 с. http://www.razym.ru/naukaobraz/psihfilosofiya/149683-bell-daniel-gryaduschee-postindustrialnoe-obschestvo-opyt-socialnogo-prognozirovaniya.html и http://www.torrentino.com/torrents/504311

7. Боярский А.Я. Статистические методы в экспериментальных медицинских исследованиях. 1955. – 262 с. Медгиз.
8. Вараксин А. Н. Статистический анализ биологической и медицинской информации: проблемы и решения / Международный журнал медицинской практики. – 2006. – № 2. – С. 35–38. http://www.biometrica.tomsk.ru/varaksin.htm
9. Власов В. В. Эпидемиология в современной России / Международный журнал медицинской практики. – 2001. – № 2. – С. 27–31.
10. Вопрос не только об автомобилях, но и об унитазах. LENTA.RU от 11 сентября 2013 г. http://lenta.ru/articles/2013/09/11/roi/
11. Воробьёв К. П. доказательная медицина – новая методология медицинской практики. ч. I. Мотивации врача и исследователя при изучении доказательной медицины / Український медичний альманах. – 2004. – № 5. – С. 41–45. http://www.vkp.dsip.net/Papers/EBM_05_L_5_1.htm
12. Воробьёв К. П. Доказательная медицина – новая методология медицинской практики. ч. II. Сущность доказательной медицины /Український медичний альманах. – 2005. – № 6. – С. 142–146. http://www.vkp.dsip.net/Papers/EBM_05_L_5_2.htm
13. Воробьёв К. П. Научные исследования и доказательная медицина: взаимоотношения и место в процессе получения новых знаний /Біль, знеболювання і інтенсивна терапія. – 2002. – № 2. – С. 2–5. http://www.vkp.dsip.net/Papers/EBM_N_Kl.htm
14. Воробьёв К. П. Проблемные вопросы представления материалов клинических исследований / Біль, знеболювання і інтенсивна терапія. – 2005. – № 2. – С. 71–80. http://www.vkp.dsip.net/Papers/Nauca_Publ.htm
15. Гор А. Земля на чаше весов. В поисках новой общей цели /Новая постиндустриальная волна на Западе. Антология. Под редакцией В. Л. Иноземцева. М.: Academia, 1999. 640 стр. с. 559. http://iir-mp.narod.ru/books/inozemcev/index.html
16. Гржибовский А. М. Корреляционный анализ в медицинских исследованиях / Бюллетень СГМУ. – 2000. – № 2. – С. 22–23.
17. Гржибовский А.М. Применение статистики в терапии: критический анализ публикаций / Бюллетень СГму. – 2000. – № 2. – С. 21–22.
18. Гржибовский А. М. Использование статистики в российской биомедицинской литературе. Экология человека . – 2008. №12, с. 55-64 http://www.biometrica.tomsk.ru/Isp_Stat.pdf
19. Громов Г.Р. Национальные информационные ресурсы: проблемы промышленной эксплуатации. Москва. Издательство «Наука», 1984. – 237 с. http://www.razym.ru/tehnicheskaya/promishlennost/233883-gromov-gr-nacionalnye-informacionnye-resursy-problemy-promyshlennoy-ekspluatacii.html
20. Гуров А.Н., Андреева И.Л. Применение информационной системы для оценки эффективности внедрения достижений медицинской науки в работу практического здравоохранения. Информационно-измерительные и управляющие системы, №12, т.8, 2010 г., стр. 47-53. http://do.gendocs.ru/docs/index-180061.html
21. Денисов Ю.Д. Информационные ресурсы в японской экономике. М.: Наука, 1994. – 208 с.

22. Дюк В. А. Преподавание прикладной статистики в Санкт-Петербургской медицинской академии последипломного образования / В. А. Дюк // международный журнал медицинской практики. – 2006. – № 2. – С. 30–31. http://www.biometrica.tomsk.ru/djuk.htm
23. ЕРЕВАНСКАЯ ДЕКЛАРАЦИЯ. http://www.biometrica.tomsk.ru/Yerevan_declaration_final_rus.pdf
24. Зорин Н. А. о неправильном употреблении термина «достоверность» в российских научных психиатрических и общемедицинских статьях /Биометрика: журнал для медиков и биологов, сторонников доказательной медицины http://www.biometrica.tomsk.ru/let1.htm . – 2000.
25. Комарова М. В. Статистический анализ биомедицинских данных: проблемы и пути решения / Международный журнал медицинской практики. – 2006. – № 2. – С. 28–29. http://www.biometrica.tomsk.ru/komarova.htm
26. Ланг Т. Двадцать ошибок статистического анализа, которые вы сами можете обнаружить в биомедицинских статьях /Международный журнал медицинской практики. – 2005. – № 1. – С. 21–31.
27. Левандовский А. Б. Доказательная медицина молодому ученому / Международный журнал медицинской практики. – 2006. – № 2. – С. 23–24. http://www.biometrica.tomsk.ru/levand.htm
28. Лем С. Сумма технологии. Изд-во «Мир», 1968. – 608 с. http://royallib.ru/book/lem_stanislav/summa_tehnologii.html
29. Леонов В. П. Долгое прощание с лысенковщиной / Биометрика. http://www.biometrica.tomsk.ru/ lis.htm
30. Леонов В. П., П. В. Ижевский Об использовании прикладной статистики при подготовке диссертационных работ по медицинским и биологическим специальностям /Бюллетень ВАК РФ. – 1997. – № 5. – С. 56–61. http://www.biometrica.tomsk.ru/leonov_vak.htm
31. Леонов В.П. Общие проблемы применения статистики в биомедицине, или что разумнее: ДДПП или ДППД? http://www.biometrica.tomsk.ru/problem.htm
32. Леонов В. П. Когда нельзя, но очень хочется, или Ещё раз о критерии Стьюдента. http://www.biometrica.tomsk.ru/student.htm
33. Леонов В. П. Обучение медиков статистике: попытка системного подхода к проблеме / в. П. Леонов // Международный журнал медицинской практики. – 2006. –№ 2. – С. 17-22. http://www.biometrica.tomsk.ru/leonov_2006.htm
34. Леонов В. П. Ошибки статистического анализа биомедицинских данных. / Международный журнал медицинской практики. – 2007. – № 2. – С. 19–35. http://www.biometrica.tomsk.ru/error_10.htm
35. Леонов В. П. Применение методов статистики в кардиологии по материалам журнала «кардиология» за 1993–1995 гг. /Кардиология. – 1998. – № 1. – С. 55–58. http://biometrica.tomsk.ru/Cardiology_1998_1.pdf
36. Леонов В. П., П. В. Ижевский Применение статистики в статьях и диссертациях по медицине и биологии. ч. 1. Описание методов статистического анализа в статьях и диссертациях / Международный журнал медицинской практики. – 1998. – № 4. – С. 7–12. http://www.biometrica.tomsk.ru/leonov_1998.htm

37. Леонов В. П. Применение статистики в статьях и диссертациях по медицине и биологии. ч. 2. История биометрики и ее применения в России / Международный журнал медицинской практики. – 1999. – № 4. – С. 7–19. http://www.biometrica.tomsk.ru/history.htm
38. Леонов В. П. Применение статистики в статьях и диссертациях по медицине и биологии. ч. 3. Проблемы взаимодействия автор – редакция – читатель / Международный журнал медицинской практики. – 1999. – № 12. – С. 7–13. http://www.biometrica.tomsk.ru/problem1.htm
39. Леонов В. П. Применение статистики в статьях и диссертациях по медицине и биологии. ч. 4. Наукометрия статистической парадигмы экспериментальной биомедицины /Международный журнал медицинской Международный журнал медицинской практики. – 2002. – № 3. – С. 6–10. http://www.biometrica.tomsk.ru/part_4.htm
40. Наукометрика статистической парадигмы экспериментальной биомедицины По материалам публикаций. Вестник Томского государственного университета. Серия "Математика. Кибернетика. Информатика". №275. Апрель 2002, стр. 17-24. http://www.biometrica.tomsk.ru/paradigma.htm
41. Леонов В.П. Доказательная или сомнительная? Медицинская наука Кузбасса: статистические аспекты. ДОКАЗАТЕЛЬНАЯ МЕДИЦИНА И СТАТИСТИКА. http://www.biometrica.tomsk.ru/kuzbass2.htm
42. Леонов В.П. Статистическая вампукизация, она же всеобщая стьюдентизация. http://www.biometrica.tomsk.ru/kuzbass5.htm
43. Леонов В.П. Современные проблемы информатики. Введение в семиотику информационных технологий: Учебное пособие. – Томск: Изд-во НТЛ, 2011. – 248 с. http://www.biometrica.tomsk.ru/lib/books/leonov_2.pdf
44. Леонов В.П. Логистическая регрессия в медицине и биологии. БИОМЕТРИКА http://www.biometrica.tomsk.ru/logit_1.htm
45. Леонов В.П. Зачем нужна статистика в доказательной медицине? Армянский медицинский реферативный журнал. 2012, вып. 9, с. 184-193. http://www.biometrica.tomsk.ru/erevan_3.html
46. Леонов В.П. Почему и как надо учить медиков статистике? Доклад на международной конференции по доказательной медицине в Ереване 18–20.10.2012. http://biometrica.tomsk.ru/erevan_4.html
47. Методы исследования операций и когнитивного анализа данных в решении задач лечебно-профилактических учреждений. Дюбанов В.В., Руднев А.С., Павловский Е.Н., Зозуля Ю.В., Самочернова А.С., Сандер Д.С. Патология кровообращения и кардиохирургия, № 4, 2011, с. 77-82. http://www.meshalkin.ru/upload/cat_mag/4_2011.pdf
48. Нижегородцев Р.М. Информационная экономика. Книга 3. Взгляд в зазеркалье. Москва — Кострома 2002. – 170 с. http://mirknig.com/knigi/estesstv_nauki/1181111404-informacionnaja_jekonomika_vse_tri_knigi.html
49. Орлов А. И. О применении статистических методов в медико-биологических исследованиях /Вестник Академии медицинских наук СССР. – 1987. –№ 2. – С. 88–94. http://www.biometrica.tomsk.ru/orlov_3.htm
50. Орчаков О.А. Теория организации. Учеб. пособие. - Изд. 2-е. – М.: МИЭМП, 2007. - 236 с.

51. Оценка медицинских технологий и формирование политики здравоохранения в странах Европы. Современное состояние, проблемы и перспективы. Marcial Velasco Garrido, Finn Borlum Kristensen, Camilla Palmhoj Nielsen, Reinhard Busse» Обсерватория Европейская по системам и политике здравоохранения. ВОЗ, 2010. – 216 с. http://www.euro.who.int/__data/assets/pdf_file/0005/128525/e91922R.pdf
52. Пильцер Пол. Безграничное богатство: Теория и практика «экономической алхимии». Новая постиндустриальная волна на Западе. Антология/Под редакцией В. Л. Иноземцева. М.: Academia, 1999. – 640 с. http://iir-mp.narod.ru/books/inozemcev/
53. Плавинский С. Л. О людях и цифрах. Обучение статистике: чему, кого и как учить? /Международный журнал медицинской практики. – 2006. – № 2. – С. 9–16. http://www.biometrica.tomsk.ru/plavin.htm
54. Производительность труда в России. http://newsruss.ru/doc/index.php/Производительность_труда_в_России
55. Рябов В. Нефтепереработка – основа стабильности экономики. Экономика и ТЭК сегодня. 2009. № 10. http://www.rusoil.ru/opinions/o505.pdf
56. Самыми неэффективными компаниями в России назвали нефтегазовые и металлургические. ЛЕТНА.РУ http://lenta.ru/news/2013/10/08/productiveness/
57. Саркисов Д.С. Некоторые особенности развития медико-биологических наук в последнее столетие //Клиническая медицина.-2000.-N7. С. 4-8.
58. Список лауреатов Нобелевской премии по физиологии или медицине http://ru.wikipedia.org/wiki/Список_лауреатов_Нобелевской_премии_по_физиологии_или_медицине
59. Счётная Палата узнала страшное: 70% госконтрактов заключают в «серой зоне». http://www.mk.ru/economics/article/2013/08/15/900166-schetnaya-palata-uznala-strashnoe-70-goskontraktov-zaklyuchayut-v-seroy-zone.html
60. Фадеев В. В., В. П. Леонов. Доказательная медицина и отечественная медицинская наука / В. В. Фадеев, В. П. Леонов, О. Ю. Реброва, Г. Л. Мельниченко // Проблемы эндокринологии. – 2003. – № 5. – С. 55–60. http://www.biometrica.tomsk.ru/naukoved/fadeev.htm
61. Федеральный справочник. Здравоохранение России. Том 13. СЕРДЕЧНО-СОСУДИСТЫЕ ЗАБОЛЕВАНИЯ ВНАЧАЛЕ XXI ВЕКА: МЕДИЦИНСКИЕ, СОЦИАЛЬНЫЕ, ДЕМОГРАФИЧЕСКИЕ АСПЕКТЫ И ПУТИ ПРОФИЛАКТИКИ. Оганов Р.Г. http://federalbook.ru/files/FSZ/soderghanie/Tom%2013/IV/Oganov.pdf
62. «Финмаркет» оценил отставание России по производительности труда. ЛЕНТА.РУ http://lenta.ru/news/2013/05/17/productivity/
63. Хрусталёв Ю.М. Философия науки и медицины: учебник. Москва.: ГЭОТАР-Медиа, – 2009. – 784 с. http://vmede.org/sait/?page=1&id=Filosofija_u4_xrustalev_2009&menu=Filosofija_u4_xrustalev_2009
64. Чернышёва Е.А. Проблемы и пути развития глубокой переработки нефти в России. Журнал «Бурение & Нефть. 2011, №5. http://burneft.ru/archive/issues/2011-05/2

65. ГОСТ Р 50779.10-2000. Статистические методы. Вероятность и основы статистики. Термины и определения. М.: Госстандарт России http://www.biometrica.tomsk.ru/GOST_50779.pdf
66. Бащинский С.Е. Статистика умеет много гитик. Международный журнал медицинской практики, 1998; №4, с.13-15 http://www.biometrica.tomsk.ru/gitiki.htm
67. Леонов В.П. ВАК для ученых? или ВАК для… бумагомарак? Троицкий вариант http://trv-science.ru/2013/04/23/vak-dlya-uchenykh-ili-vak-dlya-bumagomarak/
68. Леонов В.П. Основные понятия ROC-анализа http://www.biometrica.tomsk.ru/ROC-analysis.pdf
69. Леонов В.П. Логистическая регрессия и ROC-анализ. http://www.biometrica.tomsk.ru/logit_4.htm
70. Леонов В.П. ЛОШАДЕНДУС СВАЛЕНДУС С МОСТЕНДУС. http://www.biometrica.tomsk.ru/kuzbass6.htm
71. Доклад о человеческом развитии 2013. Опубликовано для Программы развития Организации Объединенных Наций (ПРООН) http://biometrica.tomsk.ru/PROON_2013_RU.pdf