

НАИБОЛЕЕ ЧАСТЫЕ ОШИБКИ, СОВЕРШАЕМЫЕ ПРИ ПРЕДСТАВЛЕНИИ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ

© 2020 г. А. Н. Наркевич, К. А. Виноградов

ФГБОУ ВО «Красноярский государственный медицинский университет им. проф. В. Ф. Войно-Ясенецкого», г. Красноярск

Представление результатов исследования в виде статьи, диссертации или тезисов конференции является не менее важным этапом научного исследования, чем непосредственная его реализация. Неверно представленные результаты исследования могут перечеркнуть важность и значимость даже самого методологически верно организованного и проведенного исследования. Целью данной статьи является рассмотрение наиболее частых ошибок, которые совершают исследователи на этапе публикации результатов проведенных научных медицинских исследований. Рассмотрены ошибки, связанные с неполным или избыточным описанием раздела, касающегося применяемых статистических методов, неверным выбором статистических параметров для описания медицинских данных, их ошибочным обозначением и интерпретацией, а также дефекты применения статистических критериев. Помимо этого затронуты ошибки, допускаемые при графическом представлении результатов научных исследований. Материал статьи не претендует на исчерпывающий перечень возможных ошибок, которые могут совершаться при представлении результатов исследований. Это описание опыта авторов по рецензированию и экспертизе различного рода научных публикаций и изданий. Учет изложенного в данной статье опыта позволит обеспечить более качественное приведение результатов реализованных медицинских исследований в диссертациях, статьях и материалах конференций.

Ключевые слова: методология научного исследования, ошибки в научных исследованиях, медицинские исследования, представление результатов исследования

THE MOST COMMON MISTAKES MADE BY RESEARCHERS IN PRESENTING RESEARCH RESULTS

A. N. Narkevich, K. A. Vinogradov

V. F. Voyno-Yasenetsky Krasnoyarsk State Medical University, Krasnoyarsk, Russia

Presentation of research results in the form of an article, dissertation or conference abstracts is as important stage of the project as the others. Incorrectly presented research results can devalue the importance and significance of even the most methodologically well-designed and well-conducted research. This paper reviews the most common mistakes made by researchers at the stage of publishing the results of medical research. The article presents errors related to incomplete or redundant description of statistical methods, incorrect selection of statistical parameters for describing medical data, their erroneous interpretation, as well as defects in the application of statistical criteria. In addition, errors made in the graphical representation of research results are affected. The material presented in the article does not claim to be an exhaustive list of possible mistakes that can be made when presenting research results. This material is a summary of the authors' experience in reviewing scientific manuscripts and publications. Taking into account the experience described in this article will allow for better presentation of the results of implemented medical research in theses, articles and conference materials.

Key words: research methodology, errors in scientific research, medical research, presentation of research results.

Библиографическая ссылка:

Наркевич А. Н., Виноградов К. А. Наиболее частые ошибки, совершаемые при представлении результатов исследований // Экология человека. 2020. № 8. С. 55–64.

For citing:

Narkevich A. N., Vinogradov K. A. The Most Common Mistakes Made by Researchers in Presenting Research Results. *Ekologiya cheloveka* [Human Ecology]. 2020, 8, pp. 55-64.

Представление результатов исследования в виде статьи, диссертации или тезисов конференции является не менее важным этапом научного исследования, чем непосредственная его реализация [1, 3, 4, 10]. Неверно представленные результаты исследования могут перечеркнуть важность и значимость даже самого методологически верно организованного и проведенного исследования [5, 6, 7]. Ошибки, которые допускают исследователи при представлении результатов, зачастую могут быть исправлены в процессе написания статьи или диссертации. В связи с

этим такие ошибки не носят фатального характера, но их знание может существенно упростить «жизнь» исследователю путем сокращения времени, которое потребуется на исправление и доработку его научного труда после экспертизы или рецензирования.

Данная статья не претендует на предъявление исчерпывающего списка ошибок, которые совершают исследователи при представлении результатов исследования, а лишь дает описание и обоснование целесообразности исправления наиболее частых ошибок, с которыми встречались в своей работе

авторы статьи. Цель ее — рассмотреть наиболее частые ошибки, которые совершают исследователи на этапе публикации результатов проведенных научных медицинских исследований.

Довольно частой ошибкой при оформлении статьи или диссертации является *недостаточное описание части по статистической обработке данных в разделе материалов и методов* [11, 12]. Довольно часто авторы ограничиваются выражениями типа «Использованы стандартные методы описательной и сравнительной статистики» или «Статистическая обработка данных производилась с использованием пакета прикладных программ Statistica». Этого, естественно, недостаточно по нескольким причинам. Во-первых, искушенному читателю, рецензенту или эксперту в таком случае совершенно непонятно, что означают записи, например, $5,6 \pm 2,3$ или $3,4 (1,2; 8,6)$, ведь на месте приведенных значений могут быть совершенно разные статистические параметры. Да, как правило, это среднее арифметическое \pm стандартное отклонение или медиана (первый квартиль; третий квартиль), но так ли это, читатель, исходя из скудного описания применяемых статистических методов, понять не может. Во-вторых, для искушенного читателя, рецензента или эксперта отсутствует возможность оценить, насколько верно были выбраны статистические методы и критерии и, соответственно, можно ли доверять описанным результатам, заключениям и выводам. И в-третьих, не секрет, что каждая отдельно взятая статья (речь идет именно о научных статьях) не несет существенного вклада в науку (за редким исключением). Существенный вклад вносит совокупность статей, опубликованных различными авторами, и представляющих результаты отдельных исследований по одному и тому же вопросу. Для аккумуляции информации из такой совокупности статей в мировом научном сообществе давно применяются мета-анализ и формирование систематических обзоров (не путать с научным обзором литературы, — как правило, такие обзоры описывают лишь мысли, представленные различными авторами, а не аккумуляцию их результатов). Для того чтобы использовать в мета-анализе или систематическом обзоре результат, опубликованный автором, необходимо четкое понимание статистических параметров, приведенных в статье, а при недостаточном описании раздела, посвященного описанию используемых статистических методов, это становится невозможным.

Также довольно часто *отсутствует описание параметров, приводимых в качестве значений после знака \pm* . После знака \pm могут приводиться различные параметры, например, ошибка среднего, стандартное отклонение, доверительный интервал и др. Если автор явным образом не указывает на то, что он приводит после знака \pm , то читатель не имеет возможности осознать в полной мере результаты, приведенные в статье или диссертации.

В противоположность предыдущей ошибке довольно часто возникает и обратная ошибка — *из-*

быточное описание статистических методов исследований. Это происходит, как правило, тогда, когда автор статьи или диссертации не описывает используемые в собственной работе статистические параметры и методы, а просто копирует этот раздел или из собственной предыдущей статьи, или из сторонней статьи (возможно, своего научного руководителя или коллег), которая уже была опубликована или защищена. В таком случае возникает ситуация, в которой в разделе материалов и методов описаны различные показатели, методы или критерии, а результаты их применения не представлены. Опытный рецензент или эксперт совершенно четко понимает, что в таком случае раздел, описывающий статистические методы исследований, откуда-то скопирован, и возникает вопрос — действительно ли применялись те методы, которые там описаны? Возникновение подобных вопросов заранее создает негативное отношение к результатам, описанным в статье или диссертации, у рецензента или эксперта.

Следующие три довольно частые ошибки связаны с описанием данных в статье или диссертации. Первая из них — *повсеместное использование при описании количественных данных среднего арифметического и стандартного отклонения или ошибки среднего ($M \pm \sigma$ или $M \pm m$)*. Данные параметры могут быть использованы лишь при подчинении количественных данных закону нормального распределения. В случае неподчинения количественных данных закону нормального распределения для описания таких данных должны быть использованы непараметрические аналоги. По собственному опыту авторов, количественные данные, подчиняющиеся закону нормального распределения, в реальных исследованиях встречаются довольно редко (в 5–10 % случаев).

Вторая ошибка — *использование в качестве меры разброса количественных данных ошибки среднего ($M \pm m$), а не стандартного отклонения ($M \pm \sigma$) при подчинении данных закону нормального распределения*. Это связано с тем, что ошибка среднего (m) всегда меньше стандартного отклонения (σ), что, по мнению авторов, визуально придает результатам большую точность, но опытный исследователь или специалист в области статистики знает, что это не так, так как эти показатели характеризуют данные совершенно с разных точек зрения. Описание данных реализует информационную функцию статистики, то есть параметры, с помощью которых описываются данные, должны нести в себе информацию об имеющихся у исследователя данных, на которых было проведено исследование, и соответственно приблизительную информацию об изучаемых показателях в генеральной совокупности (естественно, если выборка случайна и репрезентативна генеральной совокупности). В таком случае применение ошибки среднего (m) для описания данных является неверным, так как данный показатель не дает информации об имеющихся у исследователя данных, а показывает степень бли-

зости среднего арифметического в выборке к среднему арифметическому в генеральной совокупности. В связи с этим при описании количественных данных, подчиняющихся закону нормального распределения, вместо ошибки среднего (m) должно использоваться стандартное отклонение (σ), за исключением частных случаев, когда показать степень близости среднего арифметического в выборке к среднему арифметическому в генеральной совокупности действительно важно (такие случаи довольно редки).

Третье, наверное, нельзя назвать ошибкой, а вернее назвать исследовательским упущением или в некотором роде «недоанализом» данных — *игнорирование доверительных интервалов в качестве параметров описательной статистики*. Основная цель искушенного читателя статьи или диссертации является не просто ознакомиться с результатами, полученными автором на некоторой выборке, но и выяснить возможность использования данных результатов в своей практике, что возможно только при понимании результатов (хотя бы приблизительных), которые могут быть получены на всей генеральной совокупности. Это связано с тем, что читатель навряд ли когда-то в своей практике столкнется с имеющейся у автора выборкой, читатель будет иметь дело с генеральной совокупностью. Дать читателю результаты с определенной степенью вероятности, которые могут быть получены на всей генеральной совокупности, можно только с использованием доверительных интервалов. Доверительные интервалы могут быть рассчитаны практически для всех статистических параметров, которые исследователь рассчитывает на выборке: для среднего арифметического, стандартного отклонения, медианы, квартилей, коэффициентов корреляции, разности между средними арифметическими или медианами двух групп, показателей отношения шансов или относительного риска, чувствительности, специфичности и точности, площади под ROC-кривой и т. д. В связи с этим рекомендуется при описании данных помимо самих статистических параметров

приводить дополнительно доверительные интервалы (как правило, 95 % доверительные интервалы).

Следующей очень частой ошибкой является *интерпретация статистической корреляционной связи как зависимости или причинно-следственной связи между параметрами*. Для иллюстрации того, к чему может привести неверная интерпретация корреляционной связи, приведем следующий пример (рис. 1).

На рис. 1 представлены данные о динамике изменения расходов США на науку, космос и технологии и числа самоубийств путем повешения, удушения, удушья с 1999 по 2009 г. Коэффициент корреляции между двумя динамическими рядами составил 0,998 ($r = 0,998$), что свидетельствует о том, что между данными показателями имеется очень сильная (практически полная или функциональная) прямая корреляционная связь. Если представить, что между данными показателями имеется какая-то зависимость или причинно-следственная связь, то можно сделать заключение о том, что для того чтобы снизить число самоубийств путем повешения, удушения, удушья, необходимо снизить расходы США на науку, космос и технологии. Пример подобран специально таким образом, чтобы даже человеку, не очень разбирающемуся в подобных аналитических выкладках, было понятно, что данное заключение, естественно, неверное и может привести к фатальным последствиям для страны.

В дополнение к неверной интерпретации корреляционной связи можно привести еще одну ошибку — *неверное обозначение коэффициентов корреляции*. Зачастую коэффициенты корреляции Спирмена или Кендалла авторами обозначаются как « r », но каждый коэффициент корреляции имеет свое собственное обозначение. Так, коэффициент корреляции Пирсона обозначается как « r », коэффициент корреляции Спирмена — как « ρ » (греческая ро), а коэффициент корреляции Кендалла — как « τ » (греческая тау).

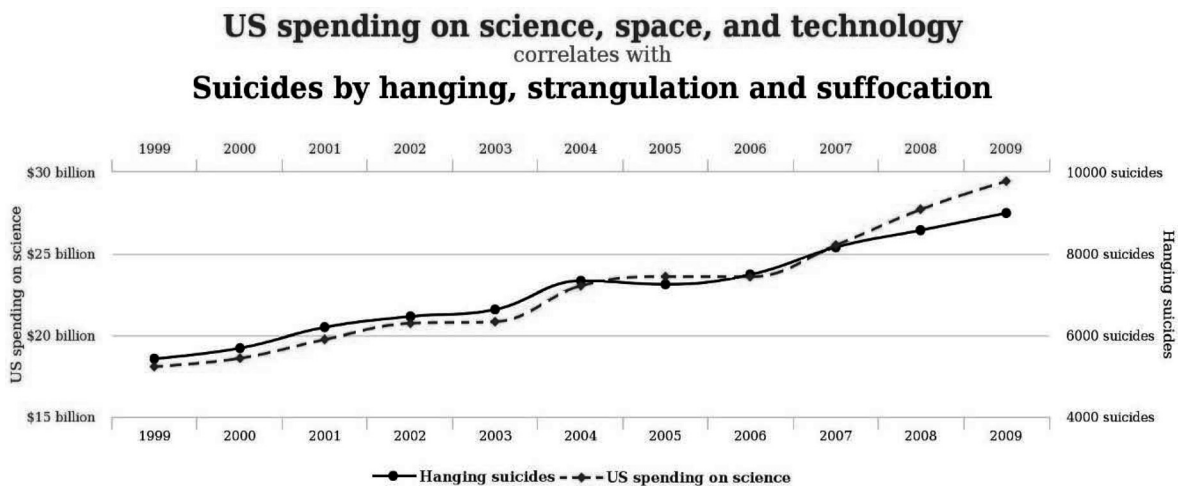


Рис. 1. Динамика расходов США на науку, космос и технологии и числа самоубийств путем повешения, удушения, удушья с 1999 по 2009 г. (<https://tylervigen.com/spurious-correlations>)

Довольно часто авторами при оценке влияния различных факторов на риск развития заболеваний или каких-либо исходов осуществляется *неправильный выбор показателя (отношения шансов или относительного риска)*. Необходимо помнить, что истинная оценка влияния фактора осуществляется с помощью показателя относительного риска, но данный показатель возможно рассчитывать лишь при проведении когортного исследования, когда группы пациентов набираются по принципу есть фактор риска/нет фактора риска. В случае проведения исследования по типу «случай-контроль», когда исследуемые группы набираются по принципу есть исход/нет исхода, приблизительная оценка влияния фактора осуществляется только с помощью показателя отношения шансов. С описанием когортных исследований и исследований по типу «случай-контроль» можно ознакомиться в одной из предыдущих статей [9].

Довольно часто исследователями *неверно используются термины «чаще» или «реже» при анализе структурных показателей (%)*. Рассмотрим две гипотетические формулировки. Первая — женщины (68,9 %) чаще болеют сахарным диабетом, чем мужчины (31,1 %). Вторая — среди женщин (68,9 %) чаще выявляется сахарный диабет, чем среди мужчин (31,1 %). На первый взгляд эти формулировки практически ничем не отличаются, но одна из них является неверной. При этом, прочитав только формулировку, сказать, какая из них неверна, нельзя. Одна из них будет неверна при определенном дизайне исследования, а другая при другом дизайне. Рассмотрим эту ошибку подробнее. Как правило, описание процентных показателей осуществляется при реализации двух классических дизайнов. Первый — проведение сравнительного исследования между двумя группами. Второй — описательное исследование одной группы. Предположим, исследователем проводилось сравнительное исследование с изучением двух групп пациентов: женщин и мужчин. Естественно, выборки сформированы с учетом требований случайности и репрезентативности. В процессе исследования установлено, что в группе женщин выявлено 68,9 % больных сахарным диабетом, а в группе мужчин — 31,1 %. В данном случае верной формулировкой заключения будет являться — среди женщин (68,9 %) чаще выявляется сахарный диабет, чем среди мужчин (31,1 %). Почему же нельзя сказать, что женщины (68,9 %) чаще болеют сахарным диабетом, чем мужчины (31,1 %)? Сам термин «чаще» (или «реже» в зависимости от результатов исследования) говорит о том, что больных женщин в абсолютном количестве больше, чем мужчин, но при сравнительном исследовании двух групп на основании процентных значений в этом нельзя быть уверенным. Так как проценты являются относительными показателями, то они не отражают никоим образом абсолютные значения больных в группах и абсолютные значения объемов этих групп. В таком случае может оказаться, что группа женщин состояла из 1 000 человек и из них

болели сахарным диабетом 689 женщин, а группа мужчин состояла из 5 000 человек и из них болели 1 555 мужчин. Если мы сделаем заключение, что женщины болеют чаще, чем мужчины, на основании процентных значений, то это окажется неверным, так как в абсолютном виде мужчин, больных сахарным диабетом, больше (1 555 человек), чем женщин (689 человек).

Исследователем проводилось описательное исследование одной группы, например, была сформирована группа больных сахарным диабетом (1 000 человек). Естественно, группа также должна быть сформирована с учетом принципов случайности и репрезентативности. В процессе исследования установлено, что среди больных сахарным диабетом женщины составили 68,9 %, а мужчины — 31,1 %. В таком случае будет верна формулировка о том, что женщины (68,9 %) чаще болеют сахарным диабетом, чем мужчины (31,1 %). В данном случае какого бы объема ни была анализируемая группа (10, 100, 1000 или 10 000 больных), 68,9 % всегда в абсолютном значении будет больше, чем 31,1 %. В связи с этим в данном случае может быть использована именно эта формулировка. Почему же нельзя сказать, что среди женщин (68,9 %) чаще выявляется сахарный диабет, чем среди мужчин (31,1 %). В данном случае несоответствие вызывает слово «среди», которое говорит о том, что раз «среди женщин», то в исследовании была группа женщин, и раз «среди мужчин», то была группа мужчин. Но в исследовании не было ни группы женщин, ни группы мужчин, а была лишь группа больных сахарным диабетом. Следует учитывать особенность использования «чаще» или «реже» при формировании заключений в статье или диссертации.

Следующий блок довольно частных ошибок при представлении результатов исследований включает в себя ошибки, связанные с представлением диаграмм. Так, *довольно часто для представления результатов выбираются неверные виды диаграмм*. Необходимо иметь в виду, что общепринятым является использование круговой диаграммы для представления структуры (в виде процентов, сумма которых равна 100) выборки или группы, линейной диаграммы (или графика) для представления изменения какого-либо показателя в динамике, столбчатые (или столбиковые) диаграммы, как горизонтальные, так и вертикальные, как правило, используются для сравнения показателей между группами, точечная — для представления связи между признаками. Естественно, в исключительных случаях эти общепринятые правила вполне оправданно могут нарушаться, как и могут применяться более современные диаграммы, такие как «ящик с усами», структурные столбиковые, лепестковые или комбинированные.

Также довольно часто встречается *отсутствие подписей осей диаграмм или отсутствие единиц измерения, в которых выражаются показатели, приводимые на осях диаграмм*. Зачастую программы, с помощью которых осуществляется создание

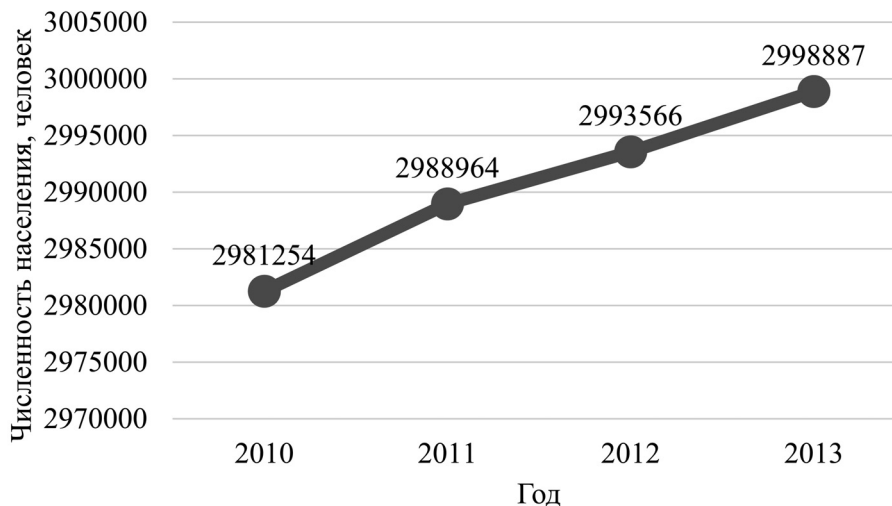


Рис. 2. Динамика численности населения Красноярского края за период с 2010 по 2013 г.

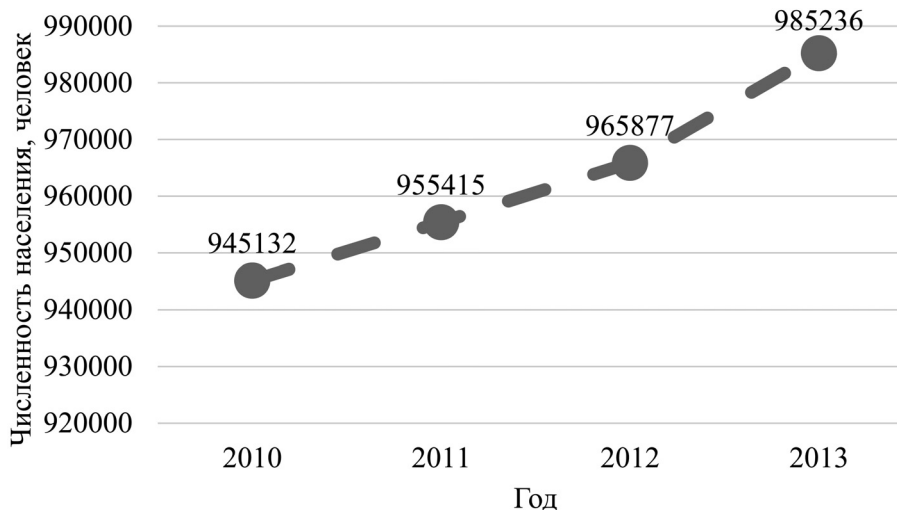


Рис. 3. Динамика численности населения города Красноярска за период с 2010 по 2013 г.

диаграмм, автоматически не вставляют эти данные в диаграмму, что вызвано различными причинами. В связи с этим необходимо помнить, что при вставке диаграммы в статью или диссертацию необходимо указать подписи ко всем имеющимся на диаграмме осям и указать единицы измерения приводимых показателей.

Также довольно частой ошибкой является *неверное представление минимальных и максимальных значений осей на диаграммах*. Довольно часто исследователи идут на это специально для того, чтобы визуально отразить существенность изменений какого-либо показателя. Так, на рис. 2 и 3 представлена динамика изменения численности населения Красноярского края и города Красноярска за период с 2010 по 2013 г. Глядя на приведенные графики, можно сказать, что за 4 года произошел существенный рост численности населения в Красноярском крае и городе Красноярске. Это связано с тем, что минимумом оси Y является довольно большое число, а приводимые на диаграмме значения слишком велики, чтобы при беглом просмотре их можно

было отнять друг от друга. Хотя даже если кому-то и удастся их бегло отнять, то получится довольно внушительное число. Так, например, в 2013 г. по сравнению с 2012 г. произошел рост численности населения Красноярского края на 5 321 человека, а в городе Красноярске – на 19 359 человек. В любой ситуации такие диаграммы создают впечатление существенной динамики показателя.

Если представить эти данные на диаграмме с минимальным значением оси Y, равным 0, то визуально данные предстанут по-другому (рис. 4). Сразу становится видно, что существенного прироста в численности населения Красноярского края и города Красноярска за 4 года не произошло. Та же самая ситуация складывается при представлении процентных значений. На рис. 5 представлено сравнение удельного веса наличия симптомов у больных двух сравниваемых групп. На первый взгляд при беглом анализе графика может показаться, что в группе I существенно ниже доля наличия сухого кашля и болей в грудной клетке, но существенно выше доля наличия такого симптома, как кровохаркание.

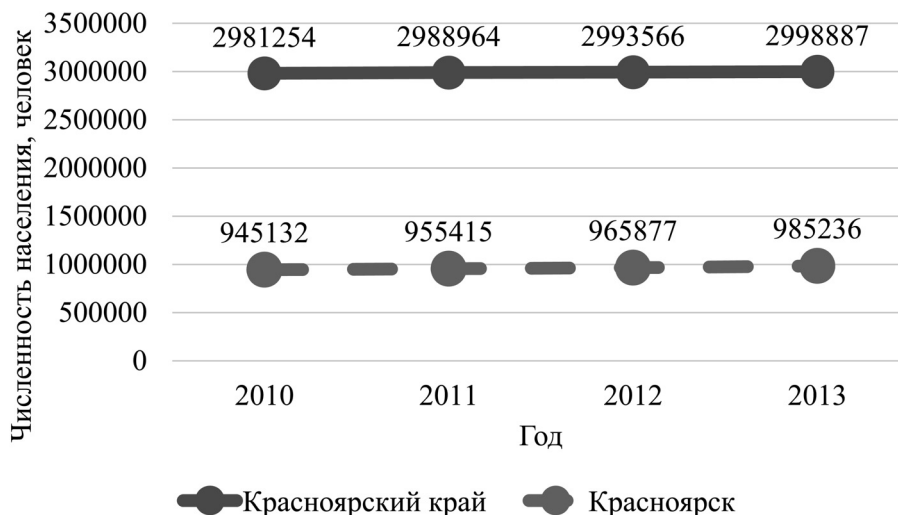


Рис. 4. Динамика численности населения Красноярского края и города Красноярска за период с 2010 по 2013 г.

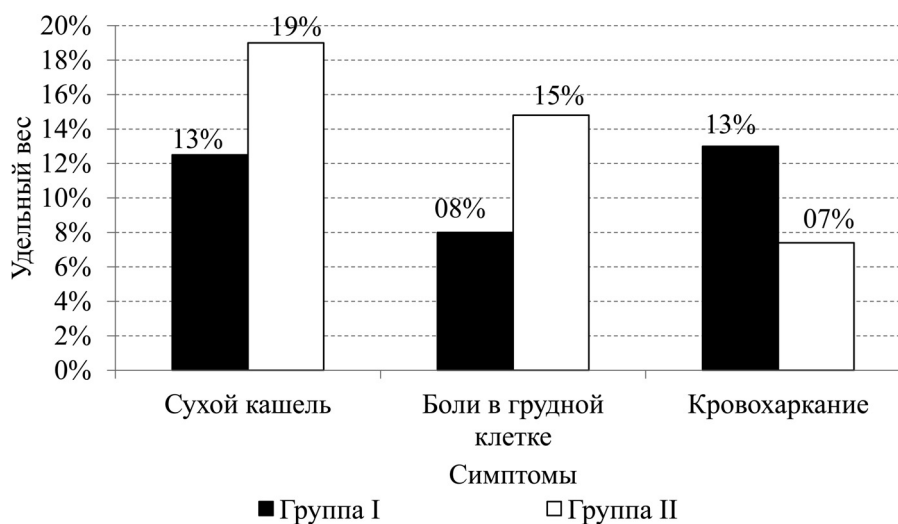


Рис. 5. Наличие симптомов у больных двух сравниваемых групп с неверным представлением максимального значения по оси Y

Однако если представить эти данные на диаграмме, на которой максимальное значение по оси Y будет равным 100 %, то визуально данные предстанут совершенно в другом виде (рис. 6). В таком случае даже при беглом анализе приведенных на диаграмме данных видно, что существенные отличия между сравниваемыми группами по наличию симптомов отсутствуют.

Следующей довольно значительной ошибкой является *интерпретация статистически значимых различий как клинически значимых*. Данная ошибка связана с широким внедрением и использованием статистических методов при проведении медицинских исследований. Ранее, когда статистические методы были менее распространены, заключение о том, что при сравнении двух исследуемых групп были установлены значительные различия уровня лейкоцитов: в 1-й группе средний уровень – $7,8 \cdot 10^9$ ед. на л, во 2-й – $8,5 \cdot 10^9$ ед. на л., в научных публикациях не возникало. Любому специалисту понятно, что разница в $0,7 \cdot 10^9$ ед. на л не имеет никакого клинического

значения, так как, во-первых, уровень лейкоцитов является довольно неспецифическим параметром, а во-вторых, такое маленькое отличие явно не имеет какого-либо практического значения. На сегодняшний же день, когда к арсеналу медицинского исследователя добавились статистические методы в дополнение к простым аналитическим заключениям, добавилось еще и статистическое подтверждение различий, и соответственно уверенность авторов в том, что действительно имеются различия между такими группами, возросла. Однако за статистическим подтверждением различий между группами зачастую теряются клинические различия. Например, исследовались две группы пациентов: 1-я группа – больные каким-либо заболеванием, 2-я группа – здоровые лица (то есть не имеющие данного заболевания). По результатам сравнения групп автором сделано заключение о том, что установлены статистически значимые различия между возрастом изучаемых групп: в 1-й группе средний возраст составил $(22,8 \pm 2,2)$ года, во 2-й

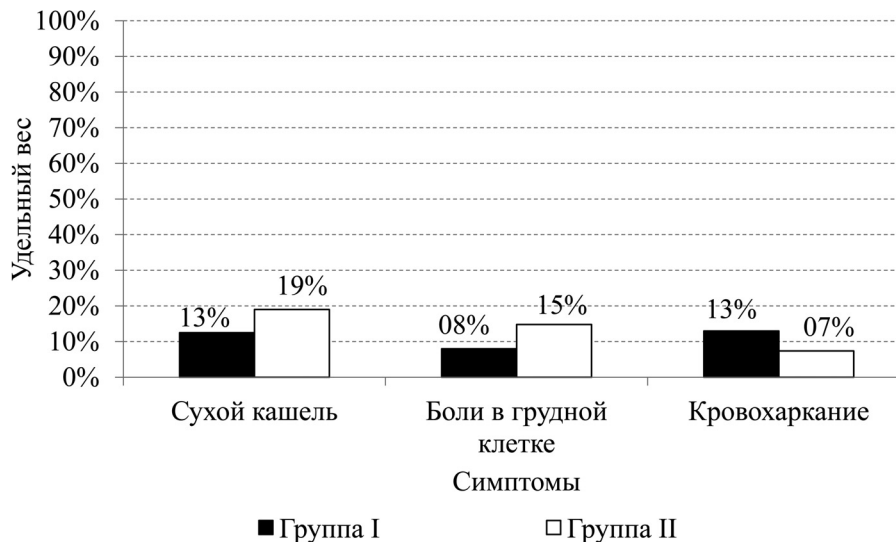


Рис. 6. Наличие симптомов у больных двух сравниваемых групп с верным представлением максимального значения по оси

группе ($27,9 \pm 2,4$) года ($p = 0,008$). Естественно, между группами по возрасту имеются статистически значимые различия, так как это подтверждено уровнем значимости примененного статистического критерия, но имеются ли между группами какие-либо клинические (то есть важные для практики) различия. Скорее нет, так как различие между группами весьма незначительное, хоть и статистически значимое. На основе такого различия практически невозможно выстраивать какую-либо клиническую тактику. Основным отличием статистической и клинической значимости является то, что статистическая значимость определяется на основании величины различия (можно назвать величиной эффекта) и объема выборок, на которых это различие получено, а клиническая значимость — только на основании величины различия. Различия может быть статистически значимо даже при минимальной величине различий, но при довольно большой выборке исследования, клинически же значимым малая величина различий быть не может. Исходя из этого, не все, что имеет статистическую значимость, является клинически значимым, но все, что является клинически значимым, должно быть и статистически значимым. Необходимо учитывать это при интерпретации результатов исследования.

Также при интерпретации результатов сравнения групп довольно часто случается *использование термина «достоверно» или «статистически достоверно»* [2]. Это возникает при описании статистически подтвержденных различий между группами. Например, между группами установлено достоверное отличие возраста ($p=0,012$). Термин «достоверно» означает 100 % уверенность в отличии, то есть все пациенты одной группы старше или младше всех пациентов второй группы. При этом такое различие сохраняется не только в изучаемых выборках, но и в генеральной совокупности в целом. Такое заключение в корне неверно. Во-первых, уровень значимости,

равный 0,012, означает, что существует вероятность того, что такие различия найдены случайно, и на самом деле в генеральной совокупности таких отличий не будет. И во-вторых, даже при очень низком уровне значимости ($p < 0,000001$) всегда данная вероятность сохраняется. В таком случае можно говорить лишь о значимом или статистически значимом различии, то есть термины «достоверно» или «статистически достоверно» необходимо заменять на «значимо» или «статистически значимо». О достоверном различии между группами, во-первых, можно говорить, когда изучается полностью вся генеральная совокупность (как правило, это результаты государственной статистической отчетности, например, заболеваемость в городе N выше, чем в городе K). Во-вторых, когда действительно все пациенты одной группы отличаются от всех пациентов второй группы (например, пациенты до 30 лет достоверно младше пациентов в возрасте 60 лет и старше). Последний пример показывает, что такие сравнения либо являются вполне очевидными, либо очень редки.

Как дополнение к предыдущей ошибке можно отметить ошибку *приведения авторами уровня значимости равного 0* ($p = 0,000$). Зачастую это связано с тем, что популярные статистические программы, в том числе IBM SPSS Statistics и StatSoft Statistica, очень маленькие значения p приводят как 0,000 или 0,00000. Такая запись указывает лишь на то, что уровень значимости очень мал ($p < 0,000001$), но не означает, что $p = 0,000$ или 0,00000. Значение уровня значимости, равное 0, означает, что различие действительно достоверное, но с помощью статистических критериев установить это невозможно. Уровень значимости всегда больше 0, хоть и может быть очень близок к 0, но все же не 0.

Еще одной схожей ошибкой является *отсутствие точного значения уровня значимости, а указание больше или меньше какого-либо граничного*

значения (как правило, 0,05). Конечно, это нельзя назвать действительно ошибкой (если правильно выбран критерий и p действительно меньше 0,05, то совершенно верно, что $p < 0,05$), скорее это можно назвать «недоописанием» результатов, которое крайне важно избегать. Если автором приводится значение уровня значимости меньше 0,05 ($p < 0,05$), то такая запись полностью лишает читателя возможности самому интерпретировать результаты исследования. Ведь $p < 0,05$ это и $p = 0,049$ и $p = 0,001$, хотя уверенность в наличии различий во втором варианте намного выше, а $p = 0,049$ может являться лишь случайностью. То же самое можно сказать и про запись о том, что $p > 0,05$ (могут быть всевозможные варианты от 0,051 до 0,547 и даже до 1,000). Заранее автор не знает, как читатель захочет воспользоваться результатом исследования. Одним для того, чтобы использовать в практике результаты исследования, будет достаточно $p = 0,049$, а для других только $p = 0,010$ и меньше. Читатели могут оценить $p = 0,051$ как намек на более тщательное изучение вопроса, освещаемого в статье или диссертации, а $p = 0,887$ как то, что в этом направлении дальнейшего изучения не требуется. Таким образом, запись $p < 0,05$ или $p > 0,05$ лишает читателя индивидуальной интерпретации результатов исследования.

Довольно частой ошибкой представления результатов исследования является *приведение в статье или диссертации 25 и 75 квартилей или 1 и 3 процентилей*. На первый взгляд может показаться, что ошибки в данном случае никакой нет, но она связана с несоответствием цифровых значений данных показателей и их названием. Так, квартили (от лат. *quarta* — четвертая) это значения в ряде данных или выборке, которые делят весь ряд на 4 равные части. То есть квартили это всего лишь 3 значения: 1 квартиль — значение, меньше которого 25 % данных и 75 % больше, 2 квартиль — значение, меньше и больше которого по 50 % данных, то есть 2 квартиль это тоже самое, что и медиана, и 3 квартиль — значение, меньше которого 75 % данных и 25 % больше. Таким образом, указание 25 и 75 квартилей попросту невозможно, так как их не существует. То же самое касается и 1 и 3 процентилей. Процентили (лат. *per cent* «на сотню; сотая») это значения, которые делят весь ряд данных или выборку на 100 равных частей. То есть существует 99 процетилей и аналогичными 1 и 3 квартилю являются 25 и 75 процентили. В связи с этим верным представлением будет являться приведение в статье или диссертации 25 и 75 процентилей или 1 и 3 квартилей.

Также довольно часто встречается *неполное или неправильное написание названий критериев*. Если автором хотя бы не полностью указано название критерия, то читатель, рецензент или эксперт не может оценить правильность его выбора в конкретном случае. Например, если в статье или диссертации указано, что применялся критерий Стьюдента, то неясно, какой — для связанных или несвязанных групп,

метод Пирсона — критерий Хи-квадрат Пирсона, или коэффициент корреляции Пирсона, или коэффициент взаимной сопряженности Пирсона, и т. д. Очень важно при описании раздела, описывающего статистические методы исследования, указывать точные и полные названия применяемых критериев. С наиболее часто применяемыми статистическими критериями при анализе медицинских данных можно ознакомиться в одной из предыдущих статей [8].

Иногда в диссертации или статье можно встретить *указание единицы измерений после медианы, а не после квартилей*. Например, медиана возраста исследуемых пациентов составила 28,5 года (20,5; 35,5), а не 28,5 (20,5; 35,5). На первый взгляд никакой особенной ошибки в этом нет и можно назвать это простой придиркой к описанию результатов. Однако приведение единицы измерения является обязательным условием приведения данных и в ситуации, когда данные представлены как 28,5 года (20,5; 35,5), ясно, в каких единицах измеряется медиана, но совершенно не ясно, в чем измеряются квартили. Запись же 28,5 (20,5; 35,5) года дает представление в каких единицах измеряется и медиана, и квартили.

Одной из довольно часто встречаемых ошибок можно упомянуть *применение статистических критериев при сравнении групп, которые представлены генеральной совокупностью*. Например, по данным государственной статистической отчетности среди мужчин больных сахарным диабетом II типа меньше, чем среди женщин ($p = 0,002$). Применение статистических критериев при сравнении групп позволяет узнать вероятность того, что имеющиеся между группами различия носят не случайный характер и будут повторяться при таком же сравнении, если заново собрать случайным образом выборки. При сравнении генеральных совокупностей (в нашем случае мужчин и женщин) отсутствует вероятность того, что имеющиеся различия установлены случайно, так как это генеральные совокупности и раз уж различия есть, то они 100 % есть. То же самое можно сказать про повторяемость такого же сравнения при новом формировании групп. Нельзя заново собрать новую генеральную совокупность, она всегда одна и та же. В связи с этим при анализе данных в генеральной совокупности отсутствует необходимость использования статистических критериев для подтверждения статистической значимости различий, так как любые различия в генеральной совокупности достоверны. Также на данных о генеральной совокупности нельзя рассчитывать доверительные интервалы. Доверительные интервалы позволяют экстраполировать результаты выборочного исследования на генеральную совокупность. То есть при расчете доверительных интервалов на данных о генеральной совокупности результаты будут экстраполироваться сами на себя.

Иногда можно встретить в статьях или диссертациях *приведение результатов статистического сравнения процентных значений в рамках одной группы*. Например, в изучаемой группе мужчин (68,5 %) было

больше, чем женщин (31,5 %; $p = 0,023$). Эта ошибка зачастую возникает с отсутствием четкого понимания групп исследования и как они сформированы или хотя бы как должны быть сформированы. Исследуемая группа формируется так, чтобы она была количественно и качественно репрезентативна генеральной совокупности, то есть группа должна представлять всю генеральную совокупность. Если это условие не соблюдено, то основываться на результатах такого исследования вообще нельзя. Исследователь сравнивает с помощью статистических критериев две правильно сформированные группы, то есть извлеченные из двух генеральных совокупностей. Например, группа мужчин и женщин — группа мужчин извлечена из генеральной совокупности мужчин, а группа женщин — соответственно из генеральной совокупности женщин. В таком случае, естественно, необходимо статистическое подтверждение различий с помощью статистических критериев. То есть необходимо подтвердить, что имеющиеся различия между группами также будут и между генеральными совокупностями, из которых извлечены данные группы. Если же исследователь сравнивает с помощью статистических критериев показатели в одной правильно сформированной группе, извлеченной из одной генеральной совокупности, то происходит сравнение данной совокупности самой с собой. Например, мужчины и женщины в рамках группы больных пневмонией пациентов — и мужчины, и женщины представляют собой одну генеральную совокупность — больных пневмонией. В таком случае если в группе больных пневмонией мужчин было больше, то и в генеральной совокупности их будет больше, так как группа сформирована случайно, она количественно и качественно репрезентативна и отражает генеральную совокупность.

Следующей ошибкой является *округление значений до различного количества знаков после запятой*. В одном случае значения представлены с округлением до двух знаков после запятой, в других — до одного, а в некоторых и вовсе до целых значений. Такое разнообразие заставляет задуматься о незнании автором правил округления. Округление структурных показателей, то есть процентов, как правило, осуществляется следующим образом. Если объем группы, из которой рассчитывается процент (значение, которое стоит в знаменателе), больше 1 000, то округление необходимо осуществлять до сотых (до двух знаков после запятой), если объем группы от 100 до 1 000 — до десятых (до одного знака после запятой), если от 20 до 100 — до целых значений процентов, а если объем группы менее 20, то вообще не рекомендуется описывать данные в виде процентов, а лучше представлять абсолютные значения. Проиллюстрировать пример абсурдности расчета процентов при малом объеме наблюдений можно следующим примером: «В результате эксперимента 33,3 % мышей выжили, 33,3 % — умерли, а третья — убежала». В отношении же не структурных показателей (то есть не процентов), как правило, при округлении исходят из следующих

условий. Во-первых, точностью прибора, которым измеряются показатели. Например, при измерении массы тела пациентов весами, погрешность которых составляет более 100 граммов, не имеет смысла округлять значения до десятых, так как при записи веса, например 63,8 кг, 8 десятых будут ошибочными. Во-вторых, необходимостью или целесообразностью округления до большого числа знаков после запятой. Например, в обычных клинических исследованиях никакого значения не имеет разница в массе тела человека 63,35 кг и 63,38 кг. Естественно, существуют исследования, при которых необходима высокая точность, в том числе и таких показателей, как масса тела человека, но это единичные случаи. Помимо правильного округления необходимо соблюдать единообразное округление, то есть округление однородных показателей до одного и того же числа знаков после запятой. Даже если вы округляете в работе все значения до сотых или до десятых и получаете в расчетах целое значение (например, 35), то его необходимо представлять не просто 35, а 35,00 или 35,0.

Очень часто можно встретить *несоответствие названий показателей, которые указываются в тексте и показателей, реально приводимых автором*. Например, в материалах и методах, описывающих статистические методы исследования, автором указано, что количественные показатели приводятся в виде медианы, 1 и 3 квартилей, а в тексте статьи или диссертации встречается фраза «Средний рост составил 163,5 (160,5; 171,5) см». В таком случае возникает несоответствие — в материалах и методах написано, что представляется медиана и квартили, реально представляется медиана и квартили (это видно из формата приведенных данных), однако называется это средним, но среднее арифметическое и медиана это два разных показателя, и они не заменяют друг друга. Если автор приводит медиану и квартили, то и в тексте должно быть указано «Медиана роста составила 163,5 (160,5; 171,5) см».

Таким образом, рассмотрены наиболее частые ошибки, совершаемые исследователями при представлении результатов медицинских исследований. Учет изложенного в данной статье опыта позволит обеспечить более качественное приведение результатов реализованных медицинских исследований в диссертациях, статьях и материалах конференций.

Авторство

Наркевич А. Н. внес существенный вклад в концепцию и дизайн исследования, получение, анализ и интерпретацию данных, подготовил первый вариант статьи, окончательно утвердил присланную в редакцию рукопись; Виноградов К. А. внес существенный вклад в концепцию и дизайн исследования, получение, анализ и интерпретацию данных, существенно переработал первый вариант статьи на предмет важного интеллектуального содержания, окончательно утвердил присланную в редакцию рукопись.

Наркевич Артем Николаевич — ORCID 0000-0002-1489-5058; SPIN 9030-1493

Виноградов Константин Анатольевич — ORCID 0000-0001-6224-5618; SPIN 6924-0110

Список литературы

1. Гржибовский А. М. Использование статистики в российской биомедицинской литературе // Экология человека. 2008. № 12. С. 55–64.

2. Зорин Н. А. «Достоверность» или «статистическая значимость» — 12 лет спустя // Педиатрическая фармакология. 2011. № 5. С. 13–19.

3. Зорин Н. А. Оценка качества научных публикаций (часть I) // Медицинские технологии. Оценка и выбор. 2011. № 3. С. 71–76.

4. Зорин Н. А. Оценка качества научных публикаций (часть II) // Медицинские технологии. Оценка и выбор. 2012. № 1. С. 85–93.

5. Ланг Т. Двадцать ошибок статистического анализа, которые вы сами можете обнаружить в биомедицинских статьях // Международный журнал медицинской практики. 2005. № 1. С. 21–31.

6. Леонов В. П. Ошибки статистического анализа биомедицинских данных // Международный журнал медицинской практики. 2007. № 2. С. 19.

7. Леонов В. П. Статистика в кардиологии. 15 лет спустя // Медицинские технологии. Оценка и выбор. 2014. № 1. С. 17–28.

8. Наркевич А. Н., Виноградов К. А. Выбор метода для статистического анализа медицинских данных и способа графического представления результатов // Социальные аспекты здоровья населения. 2019. № 4. URL: <http://vestnik.mednet.ru/content/view/1092/30/lang,ru/> DOI: 10.21045/2071-5021-2019-65-4-9.

9. Наркевич А. Н., Виноградов К. А. Дизайн медицинского исследования // Социальные аспекты здоровья населения. 2019. № 5. URL: <http://vestnik.mednet.ru/content/view/1108/30/lang,ru/> DOI: 10.21045/2071-5021-2019-65-5-13.

10. Наркевич А. Н., Виноградов К. А. Настольная книга автора медицинской диссертации: пособие. М.: Инфра-М, 2019. 454 с.

11. Реброва О. Ю. Описание статистического анализа данных в оригинальных статьях. Типичные ошибки // Российская ринология. 2018. № 1. С. 65–68.

12. Унгурияну Т. Н., Гржибовский А. М. Краткие рекомендации по описанию, статистическому анализу и представлению данных в научных публикациях // Экология человека. 2011. № 5. С. 55–60.

References

1. Grjibovski A. M. Use of statistics in Russian biomedical literature. *Ekologiya cheloveka* [Human Ecology]. 2008, 12, pp. 55-64. [In Russian]

2. Zorin N. A. «Reliability» or «statistical significance» - 12 years later. *Pediatricheskaya farmakologiya* [Pediatric pharmacology]. 2011, 5, pp. 13-19. [In Russian]

3. Zorin N. A. Assessment of the quality of scientific publications (part I). *Meditsinskie tekhnologii. Otsenka i vybor* [Medical technologies. Evaluation and selection]. 2011, 3, pp. 71-76. [In Russian]

4. Zorin N. A. Assessment of the quality of scientific publications (part II). *Meditsinskie tekhnologii. Otsenka i vybor* [Medical technologies. Evaluation and selection]. 2012, 1, pp. 85-93. [In Russian]

5. Lang T. Twenty errors of statistical analysis that you can find yourself in biomedical articles. *Mezhdunarodnyi zhurnal meditsinskoi praktiki* [International journal of medical practice]. 2005, 1, pp. 21-31. [In Russian]

6. Leonov V. P. Errors of statistical analysis of biomedical data. *Mezhdunarodnyi zhurnal meditsinskoi praktiki* [International journal of medical practice]. 2007, 2, p. 19. [In Russian]

7. Leonov V. P. Statistics in cardiology. 15 years later. *Meditsinskie tekhnologii. Otsenka i vybor* [Medical technologies. Evaluation and selection]. 2014, 1, pp. 17-28. [In Russian]

8. Narkevich A. N., Vinogradov K. A. Choice of method for statistical analysis of medical data and method of graphical representation of results. *Sotsial'nye aspekty zdorov'ya naseleniya* [Social aspects of population health]. 2019, 4, URL: <http://vestnik.mednet.ru/content/view/1092/30/lang,ru/> DOI: 10.21045/2071-5021-2019-65-4-9. [In Russian]

9. Narkevich A. N., Vinogradov K. A. Design of medical research. *Sotsial'nye aspekty zdorov'ya naseleniya* [Social aspects of population health]. 2019, 5, URL: <http://vestnik.mednet.ru/content/view/1108/30/lang,ru/> DOI: 10.21045/2071-5021-2019-65-5-13. [In Russian]

10. Narkevich A. N., Vinogradov K. A. *Nastol'naya kniga avtora meditsinskoi dissertatsii: posobie* [Table book of the author of the medical dissertation: a manual]. Moscow, Infra-M Publ., 2019, 454 p.

11. Rebrova O. Yu. Statistical analysis of the data in the original articles. Typical errors. *Rossiiskaya rinologiya* [Russian rhinology]. 2018, 1, pp. 65-68. [In Russian]

12. Unguryanu T. N., Grjibovski A. M. Brief recommendations for describing, statistical analysis, and presenting data in scientific publications. *Ekologiya cheloveka* [Human Ecology]. 2011, 5, pp. 55-60. [In Russian]

Контактная информация:

Наркевич Артем Николаевич — кандидат медицинских наук, доцент, декан-медико-психолого-фармацевтического факультета, зав. научно-исследовательской лабораторией медицинской кибернетики и управления в здравоохранении, доцент кафедры медицинской кибернетики и информатики ФГБОУ ВО «Красноярский государственный медицинский университет им. проф. В. Ф. Войно-Ясенецкого» Министерства здравоохранения Российской Федерации

Адрес: г. Красноярск, ул. Партизана Железняка, д. 1

E-mail: narkevichart@gmail.com