

От редакции: «Национальный журнал глаукома» имеет практику публикации работ, носящих спорный характер, цель которой — привлечь внимание читателя к обсуждаемой проблеме. В связи с этим мы сочли целесообразным опубликовать отзыв В.А. Аксёнова на статью А.Д. Чупрова с соавт. «Сравнительная характеристика тонометрии по Маклакову и автоматизированных методов измерения внутриглазного давления».

ОТЗЫВ

НА СТАТЬЮ ЧУПРОВА А.Д., ПИДОДНИЙ Е.А., ТРУБНИКОВА В.А.

«СРАВНИТЕЛЬНАЯ ХАРАКТЕРИСТИКА ТОНОМЕТРИИ ПО МАКЛАКОВУ И АВТОМАТИЗИРОВАННЫХ МЕТОДОВ ИЗМЕРЕНИЯ ВНУТРИГЛАЗНОГО ДАВЛЕНИЯ»,
опубликованной в «НАЦИОНАЛЬНОМ ЖУРНАЛЕ ГЛАУКОМА». 2023; 22(4):27-32.

Статья содержит серьезные ошибки, ставящие под сомнение ее научную ценность. Наиболее критическими из них являются ошибочный выбор и неправильное применение статистических методов.

Так, применявшийся в этом исследовании ранговый анализ вариаций по Краскелу–Уоллису применяется для сравнения трех и более независимых групп по одному количественному признаку. Однако в этом исследовании описываются сравнения зависимых групп. Группы являются зависимыми (связанными), когда одни и те же объекты оцениваются в разные моменты времени или разными методами измерения¹. Соответственно, группы являются зависимыми (связанными), если ВГД одних и тех же пациентов измеряется с помощью нескольких разных устройств.

Для сравнения трёх и более связанных (зависимых) групп применяется ANOVA повторных измерений (параметрический метод) или однофакторный дисперсионный анализ Фридмана (непараметрический метод)¹.

Однако эти методы не отвечают цели исследования, т.к. не позволяют оценить согласованность результатов, полученных с помощью разных методов диагностики. Т.е. они не отвечают на вопрос — дают ли два метода, используемые для измерения определенной переменной в идентичных условиях, одинаковый результат?

Выбор метода для анализа согласованности результатов, полученных с помощью разных методов диагностики, зависит от вида признаков (качественные или количественные). Для количественных переменных (например, мм рт.ст.) применяется метод Бланда–Альтмана — графический метод определения степени согласованности между двумя методами измерения одного и того же объекта².

Соответственно, метод Бланда–Альтмана применяется в исследованиях, посвященных сравнению разных методов измерения ВГД, как это можно видеть в источнике [2] (он же [8]) из списка литературы этой статьи, а также в других аналогичных исследованиях, например^{3,4,5} и др.

Другие ошибки статистического анализа:

1. Неправильно выбран формат описания переменных с нормальным распределением ($M \pm \sigma$). В статистике принято обозначать греческими буквами генеральные, т.е. популяционные параметры, тогда как выборочные (групповые) 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 — « SD ». При описании выборки эти показатели рекомендуется

1 How to Report Statistics in Medicine: Annotated Guidelines for Authors, Editors, and Reviewers. 2nd edition. Thomas A Lang and Michelle Secic. Philadelphia: American College of Physicians. 2006. 490 pages.
2 Altman DG and Bland JM. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1: 307-310.
3 Bang, S.P., Lee, C.E. & Kim, Y.C. Comparison of intraocular pressure as measured by three different non-contact tonometers and goldmann applanation tonometer for non-glaucomatous subjects. BMC Ophthalmol 17, 199 (2017). <https://doi.org/10.1186/s12886-017-0593-1>

4 Yildiz, M. B. (2023). Agreement among Goldmann Applanation Tonometer, Easyton Transpalpebral Tonometer, Tonopen, and Icare in Patients with Keratoconus. In Beyoglu Eye Journal (pp. 170–176). Kare Publishing. <https://doi.org/10.14744/bej.2023.56933>
5 Chen M, Zhang L, Xu J, Chen X, Gu Y, Ren Y, Wang K. Comparability of three intraocular pressure measurement: iCare pro rebound, non-contact and Goldmann applanation tonometry in different IOP group. BMC Ophthalmol. 2019 Nov 14;19(1):225. doi: 10.1186/s12886-019-1236-5. PMID: 31726999; PMCID: PMC6857285.

представлять в формате $M(SD)$, а не $M \pm SD$. Медиану и интерквартильный интервал рекомендуется описывать в формате $Me(Q1; Q3)$ ⁶.

2. На протяжении всего текста статьи термин «медианное значение» ошибочно заменяется на «среднее значение» (например, «среднее значение показателя составило 19 (17–20) мм рт.ст.»). Медиана и среднее арифметическое — это не одно и то же.

3. В тексте некорректно указаны качественные признаки (например, «в 7,8% случаев была диагностирована глаукома»). Абсолютные и относительные частоты необходимо указывать полностью, например, 68/1416 (4,8%).

4. Неудовлетворительно представлены результаты корреляционного анализа — не указана статистическая значимость выявленных связей.

5. Отсутствуют поправки на множественные сравнения. Поправка на множественные сравнения выполняется либо с помощью апостериорного (post-hoc) анализа, либо с помощью поправки Бонферрони — перерасчете уровня значимости p для множественных парных сравнений. Эти методы применяются при выявлении значимых различиях для уточнения того, между какими именно группами имеются различия.

6. Различия в ВГД при измерении различными методами тонометрии следовало описывать с указанием 95% доверительного интервала как показателя точности оценки. Для описания разности медиан с 95%ДИ используется оценка Ходжеса–Лемана (Hodge–Lehmann median difference⁷).

7. В тексте статьи термин «статистическая значимость» некорректно заменяется на «достоверность/достоверно»⁸. Статистические методы

не определяют «достоверность» или «недостоверность» различия, они показывают лишь вероятность случайности получения такого результата. Статистическая значимость — это малая вероятность случайности (например, выявленных различий), в то время как в теории вероятности, на которой основана математическая статистика, достоверным называется событие, вероятность которого равна 100% или 0%.

Следует также заметить, что цель исследования не соответствует содержанию статьи, т.к. цель «Изучить факторы, влияющие...» предполагает выявление факторов, которые причинно связаны с результатами измерения ВГД. Однако дизайн этого исследования и применяемые в нем статистические методы не позволяют выявлять причинно-следственную связь⁹. В данном случае можно говорить только о наличии или отсутствии статистически значимых различий и корреляций.

Заключение

В результате ошибочного выбора и неправильного применения статистических методов результаты этого исследования с большой вероятностью искажают истинные сравнительные характеристики методов измерения внутриглазного давления, что ставит под сомнение научную ценность этой статьи.

В.А. Аксёнов

12 августа 2024 г.

Аксёнов Валерий Алексеевич, канд. мед. наук
г. Оренбург
e-mail: v_aksenov@bk.ru

6 Ланг Т., Альтман Д. Основы описания статистического анализа в статьях, публикуемых в биомедицинских журналах. Руководство «Статистический анализ и методы в публикуемой литературе (САМПЛ)». Медицинские технологии. Оценка и выбор. 2014; 1(15), 11-16.

7 Staffa, S., & Zurakowski, D. (2019). Calculation of Confidence Intervals for Differences in Medians Between Groups and Comparison of Methods. Anesthesia & Analgesia. <https://doi.org/10.1213/ANE.0000000000004535>.

8 Зорин Н.А. «Достоверность» или «статистическая значимость» — 12 лет спустя. Медицинские технологии: оценка и выбор. 2011, № 2, с. 89-97.

9 Gianicolo EAL, Eichler M, Muensterer O, et al. Methods for Evaluating Causality in Observational Studies. Dtsch Arztebl Int. 2020 Feb 14; 116(7):101-107. doi: 10.3238/arztebl.2020.0101. PMID: 32164822; PMCID: PMC7081045