

УДК 615.001.31

## ПРОБЛЕМИ ОПИСУВАННЯ ТА ВИКОРИСТАННЯ СТАТИСТИЧНИХ МЕТОДІВ У НАУКОВИХ РОБОТАХ І СТАТТЯХ

О.В. Гойко

*Національна медична академія післядипломної освіти імені П.Л. Шупика*

В статті розглядаються проблеми, пов'язані з типовими помилками щодо використання статистичних методів у наукових роботах і статтях, їх інтерпретації та описуванні. Даються деякі пояснення щодо застосування найпоширеніших методів статистичного аналізу в біомедичних дослідженнях.

**Ключові слова:** статистичні методи, статистична значимість, t-критерій Стьюдента, біомедичні дослідження.

## ПРОБЛЕМЫ ОПИСАНИЯ И ИСПОЛЬЗОВАНИЯ СТАТИСТИЧЕСКИХ МЕТОДОВ В НАУЧНЫХ РАБОТАХ И СТАТЬЯХ

О.В. Гойко

*Национальная медицинская академия последипломного образования имени П.Л. Шупика*

В статье рассматриваются проблемы, связанные с типичными ошибками относительно использования статистических методов в научных работах и статьях, их интерпретации и описания. Даются некоторые объяснения относительно применения наиболее распространенных методов статистического анализа в биомедицинских исследованиях.

**Ключевые слова:** статистические методы, статистическая значимость, t-критерий Стьюдента, биомедицинские исследования.

## PROBLEMS OF THE APPLICATION OF STATISTICAL METHODS IN THE SCIENTIFIC PAPERS

O.V. Goyko

*National Medical Academy of Post-Graduate Education named after P.L. Shupyk*

In the article the common mistakes in the application of statistical methods in the scientific papers are discussed. Special explanations concerning application of most conventional statistical methods in biomedical studies are offered.

**Key words:** statistical methods, statistical significance, Student's t-test, biomedical studies.

**Вступ.** Сьогодні надзвичайно актуальною стала проблема наукового аналізу медичних даних. Завдяки використанню сучасних інформаційних технологій швидкими темпами зростають обсяги інформаційних масивів, що характеризують стан того чи іншого хворого, ускладнюються задачі, що потребують неочевидних рішень, знайти котрі можна лише після обробки достатньо великих інформаційних масивів. Сьогодні вже не підлягає сумніву теза щодо необхідності та доцільності використання доказової медицини та сучасних методів статистичного аналізу в біомедичних дослідженнях.

У медичній практиці і, особливо, у медичних дослідженнях часто застосовують різноманітні методи аналізу й обробки даних. Математика, зокрема ма-

тематична статистика, широко застосовується в усіх галузях, в тому числі у медицині, оскільки математичні методи дозволяють об'єктивно оцінювати кількісні результати досліджень.

З інтенсивним розвитком і використанням принципів доказової медицини (evidence based medicine) зросло і розуміння актуальності проблем, пов'язаних із низькою якістю опису статистичного матеріалу. В основі доказової медицини лежить використання опублікованих у медичній літературі досліджень, тому вона залежить від методологічної якості статей. В даний час спостерігається помітне відставання щодо застосування статистичних методів аналізу у вітчизняній медичній науці. Порівняння вітчизняних і зарубіжних публікацій показує набагато вищий рівень

використання методів статистики у наукових працях, що видаються за кордоном.

**Основна частина.** Намагання обговорити в медичній літературі проблеми, пов'язані зі статистичною обробкою матеріалу і статистичною ймовірністю, були започатковані ще в 30-х роках ХХ століття [1]. З того часу дослідники в різних галузях медицини вважають за необхідне використовувати методи статистичного аналізу для обробки вихідного матеріалу, допускаючи при цьому досить багато помилок як при їх використанні, так і при їх інтерпретації та описі. Неякісне використання статистичних методів в експериментальних біомедичних дослідженнях і, особливо, неправильна інтерпретація результатів їх застосування є дуже поширеною проблемою, яка має негативні наслідки. Проблема, до речі, недооцінена, не дивлячись на те, що більшість помилок виникає при використанні навіть найпростіших статистичних методів обробки даних.

Досить часто, на превеликий жаль, трапляються роботи, в яких аналіз обмежується лише якісним описом об'єктів і процесів, коли кількісні оцінки їхніх характеристик зводяться лише до констатації «збільшення» чи «зменшення» середніх значень окремих ознак. А якщо й наводяться їх числові значення, то досить часто зустрічаються вирази у такому вигляді:  $6,5 \pm 2,3$ , без вказання при цьому, які саме величини маються на увазі - чи то середнє значення і його стандартна похибка, чи то середнє значення і його середньоквадратичне відхилення. Щоб читач не ламав голову над тим, які ж саме величини поєднані знаком  $\pm$ , автору слід обов'язково вказувати загальноприйняті позначення цих величин. Можна зрозуміти авторів таких публікацій, які досить часто стоять перед вибором: що ж використовувати для опису вибіркового (а саме такі він аналізує) параметрів - чи позначення  $(M \pm t)$ , чи  $(M \pm \delta)$ , де  $t$  - стандартна похибка середнього значення,  $\delta$  - середньоквадратичне відхилення. В деяких рекомендаціях надається перевага першому варіанту, в інших (найчастіше в зарубіжних публікаціях) - другому. Стентон Гланц, наприклад, в своїй книзі «Медико-біологіческая статистика» [2] говорить наступне щодо стандартного (середньоквадратичного) відхилення і стандартної похибки: «Хотя разница между стандартным отклонением и стандартной ошибкой среднего очевидна, их часто путают. Большинство исследователей приводят в публикациях значение стандартной ошибки среднего, которая заведомо меньше стандартного отклонения. Авторам кажется, что в таком виде их данные внушают больше доверия. Может быть, так оно и есть, однако беда в том, что

стандартная ошибка среднего измеряет именно точность оценки среднего, но никак не разброс данных, который и интересен читателю. Мораль состоит в том, что, описывая совокупность, всегда нужно приводит значение стандартного отклонения». Проте, враховуючи, що "t" і "δ" тісно пов'язані між собою співвідношенням  $t = \delta / \sqrt{n}$  (де  $n$  - це корінь квадратний з обсягу вибірки "n"), то сперечатись з цього приводу, взагалі-то, немає сенсу. Вибір того чи іншого варіанта повинен в більшій мірі визначатися контекстом. Наприклад, якщо мова йде про порівняння групових середніх, то слід використовувати перший варіант. Якщо ж просто наводяться вибіркові характеристики окремих груп, то доцільніше використовувати другий варіант. Важливо, щоб як для першого, так і для другого варіанту завжди наводилось значення обсягу вибірки "n" і вказувалося, які ж величини поєднуються знаком  $\pm$ , і тоді, знаючи, одну із наведених величин, наприклад, "t" завжди можна вирахувати й значення "δ", або ж навпаки. Однак все одно це не звільняє автора від його обов'язку - чітко і конкретно вказувати всі позначення, що використовуються в його роботі.

Досить часто у публікаціях зустрічаються такі висловлювання, як, наприклад: «Позитивний кореляційний взаємозв'язок виявлений між ступенем вираженості гемофтальму та ІДЕ -  $r = \pm 0,78$ ; КАЕ -  $r = \pm 0,62$ ; ВКГ -  $r = \pm 0,96$ ; ВА АФТ-аз -  $r = \pm 0,74$ ; ІУГ -  $r = \pm 0,86$ ; від'ємний кореляційний взаємозв'язок виявлений між ступенем вираженості гемофтальму та показниками ВВЕС -  $r = \pm 0,80$ ». По-перше, якщо автор говорить про позитивний чи від'ємний кореляційний взаємозв'язок, то це виражається саме знаком перед наведеним коефіцієнтом кореляції, а не писати перед значенням наведеного коефіцієнта  $\pm$ . Такий знак ( $\pm$ ) ставиться перед коефіцієнтом кореляції, коли мова йде про загальне поняття коефіцієнта кореляції, вказуючи тим самим, що він може приймати як позитивне, так і від'ємне значення, і саме цей знак «+» чи «-» вказує на прямий чи зворотний кореляційний взаємозв'язок. По-друге, коли в роботі наводиться кореляційний аналіз, то обов'язково слід вказати, який саме коефіцієнт кореляції (Пірсона, Спірмена, Кендала тощо) використовувався. І нарешті, що найголовніше, не наводиться (а напевне і не визначається) коефіцієнт надійності, який вказує ступінь довіри до визначеного коефіцієнта кореляції. Адже можна і на двох показниках визначити коефіцієнт кореляції, який за своїм значенням буде досить високим, проте при визначенні коефіцієнта надійності ми впевнимся у тому, що такому коефіцієн-

ту кореляції довіряти не можна і виявлений кореляційний взаємозв'язок не можна вважати суттєвим.

Отже, автори завжди повинні наводити не лише значення коефіцієнта кореляції при використанні кореляційного аналізу, а й визначати і, звичайно, наводити в публікаціях, коефіцієнт надійності.

Ще хочеться торкнутися досить слушного питання щодо порівняння показників у двох групах (наприклад, контрольній і експериментальній). Без такого порівняння, мабуть, не можна обійтися в жодній науковій роботі. Так, наприклад, автор однієї статті стверджує, що «... тривалість психозу у хворих з аритмією збільшується з 2,75 до 3, 16 днів ( $p < 0,05$ ), з цирозом з 2,5 днів до 3,10 днів ( $p < 0,01$ )...». В чому ж тут проблема? На перший погляд ніби й нормально все, приводиться  $p$ , і так, на превеликий жаль, більшість авторів пишуть. Проте, наводячи такі показники (наскільки можна зрозуміти, що це їх середні значення), слід обов'язково вказувати ще й стандартну похибку середнього значення, про що вже йшлося вище. Адже статистично достовірну значимість різниці середніх значень кількісних показників саме у двох вибірках найчастіше визначають за допомогою параметричного Іжритерію Стьюдента, який надзвичайно популярний і використовується у більшості медичних наукових публікацій [3]. А як відомо, цей критерій обчислюється за допомогою двох параметрів - вибірових середніх значень і їх стандартної похибки. Саме тому при згадуванні середніх значень у такому контексті необхідно обов'язково вказувати і стандартну похибку, оскільки досвідчений читач саме завдяки останній може зорієнтуватися щодо правильності наведених тверджень. Разом з тим, хочеться ще раз наголосити на тому, що досить часто Іжритерій Стьюдента використовується невірно. Насамперед слід нагадати, за яких умов можна використовувати цей критерій. По-перше, згаданий критерій використовується лише для кількісних величин, що підлягають закону нормального розподілу [4]. По-друге, навіть за умови нормальності розподілу кількісних величин, слід пам'ятати, що таких Іжритеріїв існує декілька, а саме: ^критерій для незалежних вибірок (такими є, наприклад, контрольна і експериментальна групи), при чому в цьому випадку такий критерій обчислюється з урахуванням визначеної дисперсії (розраховується 1-критерій при однакових чи різних дисперсіях), і так званий парний критерій, що розраховується для парних вибірок (наприклад, порівняння показників до і після експерименту) [5]. У випадку, коли варіаційний ряд не підпорядкований нормальному закону розподілу, то 1-кри-

терій Стьюдента використовувати не можна, тоді найчастіше використовується непараметричний критерій  $\chi^2$  (хі-квадрат). Лише після коректного вибору і визначення відповідного критерію можна визначати значення досягнутого рівня статистичної значимості, яке бажано подавати у такому вигляді:  $p=0,03$ , а не вказувати його інтервальну оцінку, тобто писати  $0,01 < p < 0,05$  чи  $p > 0,05$ . До речі, рівень статистичної значимості позначається маленькою літерою  $p$ , а не великою літерою  $P$ , оскільки це зовсім протилежне поняття. Ще хочеться наголосити на тому, як наводяться ці позначки у таблицях. Досить часто у таблицях вони подаються у вигляді \* (зірочок), що стоять і зверху, і знизу, і зліва, і справа біля, наскільки можна зрозуміти, бо це ніде не вказується, середніх показників із знаком «+», а під таблицею стоїть примітка такого змісту: «\* $P < 0,02$ ;  $P^* < 0,05$ ; \* $P < 0,05$ ;  $0,1 > P^* > 0,05$ ». Розібратися у цій плутанині просто неможливо. Напевне і сам автор такого «ребусу» не зможе пояснити, чим, наприклад,  $P^* < 0,05$  відрізняється від \* $P < 0,05$ . По-перше, якщо автор мав на увазі рівень статистичної значимості, то його, як вже говорилося, слід позначати маленькою літерою  $p$ . По-друге, абсолютно незрозуміло, що і з чим порівнюється. А по-третє, уявіть собі таблицю, в якій біля одного числа стоїть три \*, одна знизу зліва, друга зверху зліва, а третя справа зверху, і при цьому, ще раз наголошую, не сказано, які ж показники порівнюються між собою. Найімовірніше, з цього приводу автор і сам не зможе нічого пояснити. Можливо так пишуть ті автори, які, на жаль, не розуміють суті наведених таблиць: «туманно пишеш про те, що погано розумієш». Слід пам'ятати, що опис використаних для обробки медичної інформації статистичних методів і отриманих в результаті їх застосування висновків дає читачеві інформацію ще й про рівень володіння автором цим методом. Саме тому кожен, хто пише наукову роботу, повинен описувати лише те, що він сам добре розуміє, і не вводити в оману читача. Досить часто автори не розуміють, що вони хочуть отримати від використання статистичних методів. Одні використовують статистичний аналіз тому, що так робить більшість їхніх колег, інші вважають, що наявність статистичних висновків підніме науковий рівень публікації, викликавши більший інтерес до неї.

**Висновок.** Бажано всім авторам публікацій коректно описувати і усвідомлено використовувати статистичні методи, щоб на основі отриманих результатів обробки й аналізу вибірових спостережень зробити обґрунтований висновок щодо всієї досліджуваної популяції. Застосування статистичних ме-

## МЕДИЧНА ІНФОРМАТИКА ТА ІНЖЕНЕРІЯ

тодів повинно впевнити читачів їхніх робіт у достовірності наукових стверджень завдяки використанню математичної статистики в біомедицині. Для цьо-

го слід надавати у наукових роботах і статтях ясний, повний і коректний опис виконаного статистичного аналізу спостережень.

### Література

1. Mainland D. Chance and the blood count. 1934 CMAJ 1993; 148:225-7.
2. Сієнтон Гланц. Медию-биологическая статистика / Гланц С. - М.: Практика, 1999. - 220 с.
3. Feinstein A.R. Clinical biostatistics: a survey of statistical procedures in general medical journals. Clin.Pharmacol. Ther., 15:97-107, 1974.
4. Мінцер О.П. Оброблення клінічних і експериментальних даних у медицині: навч. посібник / О.П. Мінцер, Ю.В. Вороненко, В.В. Власов. - К.: Вища школа, 2003. - 350 с.
5. Лапач С.Н. Статистические методы в медико-биологических исследованиях с использованием Excel / С.Н. Лапач, А.В. Чубенко, П.Н. Бабич. - К.: Морион. - 2000. - 315 с.