

*Мусатов М.И., доктор медицинских наук,
НИИ фундаментальной и клинической иммунологии*

*Настоящее исследование выполнено при финансовой поддержке
Российского гуманитарного научного фонда (проект №16-03-00345 «Наукометрическая
оценка качества отечественных исследований в клинической иммунологии»)*

ОПЫТ ПРИМЕНЕНИЯ ЕДИНОЙ КАРТЫ ДЛЯ ФОРМАЛИЗОВАННОЙ ЭКСПЕРТНОЙ ОЦЕНКИ КАЧЕСТВА НАУЧНЫХ ПУБЛИКАЦИЙ В КЛИНИЧЕСКОЙ ИММУНОЛОГИИ

Аннотация: целью настоящей работы был опыт анализа отечественных публикаций в области изучения клеточного иммунитета за 2008-2015 гг. с использованием единой карты (ЕК) как инструмента для формализации и последующего анализа информации в трех аспектах – статистической, медицинской и общенаучной доказательности.

Результаты: анализировались две группы публикаций: 1) использовавшие при анализе результатов абсолютные значения (22 статьи) и 2) использовавшие абсолютные и относительные значения как независимые переменные (44 статьи). При сравнении по 11 вопросам ЕК в 10 получены статистически значимые различия.

Вывод: использование ЕК позволяет получать информацию, пригодную для дальнейшей обработки и оценки методами биостатистики.

Ключевые слова: иммунология; библиометрия; квалиметрия

Первые зарубежные исследования, посвященные применению статистических методов в медицине, были проведены в 1929 г. и 1932 г. В СССР же первое масштабное исследование такого плана было организовано по инициативе профессора МГУ В.В. Налимова и завершено в 1990 г. [1]. В последствие специалисты по биостатистике неоднократно возвращались к анализу многочисленных ошибок, допускаемых отечественными авторами в своих публикациях [2, 3]. Важность этой проблемы трудно переоценить, ведь некорректная оценка полученных данных фактически «обнуляет» ценность результатов проведенных исследований.

Ранее мы предложили как инструмент для ква-

лиметрической оценки публикаций в клинической иммунологии «Единую карту» (ЕК), которая представляет собой модифицированный опросник, созданный Н.А. Зориным и соавторами [4].

Целью настоящей работы был опыт анализа отечественных публикаций в области изучения клеточного иммунитета с использованием единой карты как инструмента для формализации и последующего анализа в трех аспектах – статистической, медицинской и общенаучной доказательности.

Методы. Дизайн исследования. Наукометрическое ретроспективное исследование. Использовались публикации журналов «Иммунология» и «Медицинская иммунология» за 2008-2015 гг. По-

иск в «Медицинской иммунологии» осуществлялся по электронным версиям журналов, в «Иммунологии» – по полиграфической версии издания. *Критерии включения.* В анализ включались только публикации, являвшиеся структурированными статьями, где объектами исследования были здоровые и больные люди. Включались публикации, где авторы использовали проточный цитометрический анализ или метод непрямой иммунофлюоресценции. *Критерии исключения.* В анализ не включались статьи, чьи выводы основывались на исследованиях *in vitro*.

В целом из 856 статей на человеческом материале в анализ было включено 188 работ, посвященных иммунофенотипированию клеток. Из них были сформированы две группы сравнения. Первую составили 22 работы, авторы которых использовали в анализе только абсолютные значения клеточного счета (группа 1). По нашим данным, полученным ранее, это все публикации такого типа в анализируемых изданиях, за исключением возможных случайных ошибок при анализе журнальных массивов. Вторую группу составили 44 работы, чьи авторы использовали относительные и абсолютные значения содержания клеток как независимые переменные (группа 2).

В оригинале «единой карты» [4] ответы на все вопросы имеют три градации: 2 – выполнено полностью; 1 – выполнено частично (или неясно, не описано); 0 – не выполнено. Для вопросов, подразумевающих бинарную оценку, используются баллы 2 либо 0. В результате анализируемые статьи получают определенный средний балл, после чего можно сопоставлять их распределение по сравниваемым группам. В настоящей работе мы сравнили выделенные группы публикаций по долям баллов «2» в основных пунктах нашего опросника.

Такая форма представления результатов в настоящий момент представляется оправданной для демонстрации возможностей ЕК, поскольку она может конкретно показать, какие именно основные отличия есть между группами сравнения.

Статистическая обработка данных. 95% доверительный интервал для отношений определялся с помощью свободно распространяемой программы «Winperі» версия 11.5, 2015 г («Программы для эпидемиологов для Windows») [Abr.], с ее же помощью оценивалась статистическая значимость при сравнениях в таблице 2 x 2 с помощью скорректированного критерия χ^2 (Upton's «N-1» χ^2). Относительные частоты приведены в процентах.

Этические аспекты. Авторы разделяют и поддерживают точку зрения А.М. Гржибовского, заместителя главного редактора журнала «Экология человека», высказанную им в публикации, посвященной анализу ошибок в статистической обработке данных – мы знакомим «...будущих авторов с наиболее часто встречающимися ошибками, а не с авторами, которые эти ошибки допускают» [2, с. 55], поэтому ссылки на анализируемые публикации намеренно не приводятся.

Результаты. Приводим доли оценок «2», присвоенным двум группам статей при «ответе» на вопросы ЕК.

Группа основных вопросов о статистической значимости исследования

Вопрос ЕК: определение необходимого объема выборки.

В выделенных группах исследований данный вопрос не рассматривался.

Вопрос ЕК: наличие и полнота описания статистических методов анализа.

В группе 1: 64%, в группе 2: 18%. В группе 1

оценка «2» встречается чаще в 3,5 раза (ДИ 1,3 – 9,44, $p = 0,01$). В группе 2 оценку «0» получили такие описания как проведение статистической обработки «...при помощи автоматической компьютерной программы Microsoft Exel», или сообщения о загадочных конструктах типа «LSD-критерий Фишера» или «критерий Стьюдента в модификации Уэлча». В некоторых работах описание статметодов вообще отсутствует, но в тексте появляются сообщения типа « $p < 0,05$ ». В.П. Леонов называл это явление «самозародившимся р».

Вопрос ЕК: использование описательных статистик для количественных данных, нормальное распределение.

В группе 1: 20% и в группе 2 – 17% (разница статистически незначима).

Вопрос ЕК: использование описательных статистик для количественных данных, распределение, отличное от нормального.

В группе 1: 54 %, в группе 2 – 14%, т.е. в 4 раза реже (ДИ (1,73 – 9,23), $p < 0,0001$). В остальных 68% работ из группы 2 используется средняя с ее ошибкой (что поясняется, либо средняя, соединенная знаком «плюс/минус» с чем-то без пояснений).

Вопрос ЕК: анализ соответствия вида распределения признаков закону нормального распределения.

В группе 1: 55%, в группе 2: 27%, то есть в группе 1 оценка «2» встречается в 2 раза чаще (ДИ 1,1 – 3,62, $p = 0,011$). Собственно анализ «нормальности» распределения нужен для дальнейшего использования параметрических статистических критериев, а в отечественной иммунологии и биомедицине в целом из них наиболее распространен t -критерий Стьюдента.

Вопрос ЕК: использование t -критерия Стью-

дента для количественных признаков для двух независимых выборок (при условии соответствия распределения признаков нормальному закону и равенству дисперсий).

В группе 1 данный критерий используется в 26% случаев и определение равенства дисперсий также используется в 26% случаев (как правило, использовался F -критерий Фишера, реже критерий Левена). В группе 2 данный вариант критерия Стьюдента использовался в 54% случаев, но определения равенства дисперсий не проводилось. Несмотря на то, что часть исследований все-таки тестировала наличие нормальности распределения в выборках, без выполнения второго обязательного условия применимости данного варианта t -критерия 54% статей из группы 2 не могут получить иную оценку как «0».

Вопрос ЕК: использование критерия Манна-Уитни для количественных или порядковых признаков для сравнения двух независимых выборок.

В группе 1: 64%, в группе 2 – 25%, то есть в 2,33 раза реже (ДИ (1,31 – 4,15, $p = 0,005$).

Группа вопросов об общенаучной значимости исследования

Вопрос ЕК: обсуждается ли влияние случайных ошибок на результаты исследования на аналитическом (и постаналитическом) этапах исследования?

В обеих группах данные вопросы не обсуждаются.

Вопрос ЕК: описаны ли критерии включения / исключения объектов исследования?

В группе 1: 91%, в группе 2 – 54%, в 1,7 раз реже (ДИ 0,44 – 0,81, $p = 0,003$).

Вопрос ЕК: дана ли характеристика включенных больных?

В группе 1: 100%, в группе 2 – 82%, т.е. в 1,2

раза (ДИ (1,06 – 1,4), $p = 0,034$).

Вопрос ЕК: учитывались ли возможные источники систематических ошибок при планировании, анализе или обсуждении результатов?

В обеих группах данные вопросы не обсуждаются.

Вопрос ЕК: соответствуют ли формально выводы заявленной цели исследования (гипотезе)?

В группе 1: 72%, в группе 2 – 41%, т.е. 1,8 раза реже (ДИ (1,15 – 2,75), $p = 0,016$).

Вопрос ЕК: основаны ли выводы на полученных результатах?

В группе 1: 73%, в группе 2 – 18%, т.е. в 4 раза реже (ДИ (2,03 – 7,87), $p < 0,0001$).

Обсуждение. Принципы научной медицины начинают использоваться в иммунологических исследованиях, но в анализируемых группах публикаций они не применялись и поэтому их медицинская значимость не оценивалась.

Полученные данные демонстрируют более высокое качество тех исследований, где авторы использовали абсолютные значения содержания клеток. Очевидно, что это некоторым образом связано и с другими аспектами в организации научного исследования.

При использовании абсолютных и относительных значений, выводы делались по последним, если они «показали» статистически значимые различия, а абсолютные – нет, что снижает степень

обоснованности выводов полученными результатами и степень соответствия выводов заявленной цели исследования.

Выявленные недостатки статистической обработки полученных данных не демонстрируют принципиальных отличий от данных других исследований, посвященных анализу применения статистических методов в биомедицинских исследованиях [1, 2, 3]. В то же время вопросы об общенаучной значимости исследований в иммунологии не проводились. Например, вопрос об определении необходимого размера выборок до начала исследований фактически относится к компетенции биостатистики, поскольку она разработала методы для решения данного вопроса. С другой стороны, неадекватный объем обследованных лиц сам по себе, а в случае обследования больных еще и связанный с их неизбежной гетерогенностью по длительности заболевания, его активностью и т.п. является главным источником систематических ошибок, приводящих к смещенности результатов исследования.

В целом метод экспертного шкалирования, положенный в основу оценки качественных признаков в ЕК, показал возможность получения информации, позволяющей проводить ее дальнейшую обработку и оценку методами биостатистики: в 10 из 11 анализируемых вопросов получены статистически значимые различия.

Литература

1. Леонов В.П., Ижевский П.В. Применение статистики в медицине и биологии: анализ публикаций 1990-1997 гг. // Сибирский медицинский журнал. 1997. Вып. 3-4. С. 64 – 74.
2. Гржибовский А.М. Использование статистики в российской биомедицинской литературе // Экология человека. 2008. №12. С. 55 – 64.
3. Леонов В.П. Статистика в кардиологии. 15 лет спустя // Медицинские технологии. Оценка и выбор. 2014. №1 (15). URL: http://www.biometrica.tomsk.ru/stat_cardio1.htm. (дата обращения: 10.12.2017)

4. Зорин Н.А., Немцов А.В., Калинин В.В. Формализованная экспертная оценка качества исследовательских публикаций в психиатрии // Рекомендации по подготовке научных медицинских публикаций: Сборник статей и документов / под ред. С.Е. Бацинского, В.В. Власова. М., Медиа Сфера, 2006. С. 442 – 459.

5. Abramson J.H. WINPEPI updated: computer programs for epidemiologists, and their teaching potential // Epidemiol. Perspect. Innov. 2011. 8 (1):1. doi: 10.1186/1742-5573-8-1.

References

1. Leonov V.P., Izhevskij P.V. Primenenie statistiki v medicine i biologii: analiz publikacij 1990-1997 gg. // Sibirskij medicinskij zhurnal. 1997. Vyp. 3-4. S. 64 – 74.

2. Grzhibovskij A.M. Ispol'zovanie statistiki v rossijskoj biomedicinskoj literature // EHkologiya cheloveka. 2008. №12. С. 55 – 64.

3. Leonov V.P. Statistika v kardiologii. 15 let spustya // Medicinskie tekhnologii. Ocenka i vybor. 2014. №1 (15). URL: http://www.biometrica.tomsk.ru/stat_cardio1.htm. (data obrashcheniya: 10.12.2017)

4. Zorin N.A., Nemcov A.V., Kalinin V.V. Formalizovannaya ehkspertnaya ocenka kachestva issledovatel'skih publikacij v psihiatrii // Rekomendacii po podgotovke nauchnyh medicinskih publikacij: Sbornik statej i dokumentov / pod red. S.E. Bashchinskogo, V.V. Vlasova. М., Медиа Сфера, 2006. С. 442 – 459.

5. Abramson J.H. WINPEPI updated: computer programs for epidemiologists, and their teaching potential // Epidemiol. Perspect. Innov. 2011. 8 (1):1. doi: 10.1186/1742-5573-8-1.

*Musatov M.I., Doctor of Medical Sciences (Advanced Doctor),
Federal Scientific Institute of Fundamental and Clinical Immunology*

THE EXPERIENCE OF USING A UNITED CARD FOR A FORMAL EXPERT ASSESSMENT OF THE QUALITY OF SCIENTIFIC PUBLICATIONS IN CLINICAL IMMUNOLOGY

Abstract: the aim of this work was the analysis of the publications in the study of cellular immunity during 2008-2015, using a united card (UC) as a tool for the formalization and subsequent analysis of information in three ways – statistical, medical and scientific evidence.

Results: two groups of publications were analyzed: 1) used in the analysis of the results absolute values of cell count (22 articles) and 2) used in the analysis of the results absolute values of cell count and relative (percentage) values as independent variables (44 items). When comparing the 11 questions received in the UC 10 statistically significant differences.

Conclusion: the use of UC allows obtaining information suitable for further processing and evaluation methods of biostatistics.

Keywords: immunology; bibliometrics; qualimetry