

22. Bradley C. // New Trends in Patient Education: A Trans-Cultural and Inter-Disease Approach. International Congress Series 1076 / Eds J.-Ph. Assal et al. — New York, 1995. — P. 99—106.
23. Campbell E. M., Redman S., Moffitt P. S., Sanson-Fisher R. W. // Diabetes Educ. — 1996. — Vol. 22. — P. 379—386.
24. Clement S. // Diabetes Care. — 1995. — Vol. 18, N 8. — P. 1204—1214.
25. Cook C., Lyles R., El-Kebbi I. et al. // Ibid. — 2001. — Vol. 24. — P. 209—215.
26. Delamater A. M., Jacobson A. M., Anderson B. et al. // Ibid. — N 7. — P. 1286—1292.
27. Frost G., Wilding J., Beecham J. // Diabet. Med. — 1994. — Vol. 11. — P. 397—401.
28. Funnell M., Anderson R., Arnold M. et al. // Diabetes Educ. — 1991. — Vol. 17. — P. 37—41.
29. Glasgow R., Osteen V. // Diabetes Care. — 1992. — Vol. 15, N 10. — P. 1423—1432.
30. Glasgow R., Anderson R. // Ibid. — 1999. — Vol. 22, N 12. — P. 2090—2091.
31. Grey M., Boland E. A., Davidson M. et al. // Appl. Nurs Res. — 1999. — Vol. 12. — P. 3—12.
32. Jacobson A. M. // N. Engl. J. Med. — 1996. — Vol. 334, N 19. — P. 1249—1253.
33. Laitinen J., Ahola I., Sarkkinen E. et al. // J. Am. Diet. Assoc. — 1993. — Vol. 93. — P. 276—283.
34. Lane J. D. // Diabetes Care. — 2000. — Vol. 23, N 9. — P. 1321—1325.
35. Litzelman D. K., Slemenda C. W., Langefeld C. D. et al. // Ann. Intern. Med. — 1993. — Vol. 119. — P. 36—41.
36. Lorenz R. A. // Diabetes Educ. — 1989. — Vol. 15. — P. 149—152.
37. Lorenz R. A., Wysocki T. // Diabetes Spectrum. — 1991. — Vol. 4. — P. 261—292.
38. Lustman P. J., Griffith L. S., Freedland K. E. et al. // Ann. Intern. Med. — 1998. — Vol. 129. — P. 613—621.
39. Maldonado A., Bloise D., Ceci M. et al. // Patient Educ. Couns. — 1995. — Vol. 26. — P. 57—66.
40. Meltzer L. J., Johnson S. B., Prine J. M. et al. // Diabetes Care. — 2001. — Vol. 24. — P. 678—682.
41. Mensing C., Boucher J., Cypress M. et al. // Ibid. — 2002. — Vol. 25. — P. S140—S147.
42. Miller L., Goldstein J. // N. Engl. J. Med. — 1972. — Vol. 286. — P. 1388—1391.
43. Miller-Johnson S., Emery R. E., Marvin R. S. et al. // J. Consult. Clin. Psychol. — 1994. — Vol. 62. — P. 603—610.
44. Mollema E. D., Snoek F. J., Ader H. J. et al. // J. Psychosom. Res. — 2001. — Vol. 51, N 5. — P. 665—672.
45. Muhlhauser I., Bruckner I., Berger M. et al. // Diabetologia. — 1987. — Vol. 30. — P. 681—690.
46. Noel P. H., Larne A. C., Meyer J. et al. // Diabetes Care. — 1998. — Vol. 21, N 6. — P. 896—901.
47. Norris S. L., Engelgau M. M., Narayan K. M. V. // Ibid. — 2001. — Vol. 24, N 3. — P. 561—587.
48. O'Connor P. J., Crabtree B. F., Abourizk N. N. // J. Am. Board Fam. Pract. — 1992. — Vol. 5, N 4. — P. 381—387.
49. Perry T., Mann J., Lewis-Barned N. et al. // Eur. J. Clin. Nutr. — 1997. — Vol. 51. — P. 757—763.
50. Peterson K. A., Hughes M. // J. Am. Board Fam. Pract. — 2002. — Vol. 15, N 4. — P. 266—271.
51. Prochaska J. A. // Cancer. — 1991. — Vol. 67. — P. 805—807.
52. Reaven G. M., Thompson L. W., Nahum D., Haskins E. // Diabetes Care. — 1990. — Vol. 13. — P. 16—21.
53. Reynaert C., Janne P., Donckier J. et al. // Diabete Metab. — 1995. — Vol. 21, N 3. — P. 180—187.
54. Rosilio M., Cotton J. B., Wieliczko M. C. et al. // Diabetes Care. — 1998. — Vol. 21, N 7. — P. 1146—1153.
55. Rubin R. R., Peyrot M., Saudek C. D. // Ibid. — 1989. — Vol. 12, N 10. — P. 673—679.
56. Sakharova J. V., Yefimov A. S. // Diabetologia. — 2002. — Vol. 45. — Suppl. 2. — P. A20.
57. Stott N. C. H., Rees M., Rollnick S. et al. // Patient. Educ. Couns. — 1996. — Vol. 29. — P. 67—73.
58. Surwit R. S., Tilburg M., Zucker N. et al. // Diabetes Care. — 2002. — Vol. 25, N 1. — P. 30—34.
59. The Diabetes Control and Complications Trial Research Group // N. Engl. J. Med. — 1993. — Vol. 329. — P. 977—986.
60. Trief P. M., Grant W., Elbert K., Weinstock R. S. // Diabetes Care. — 1998. — Vol. 21, N 2. — P. 241—245.
61. UK Prospective Diabetes Study (UKPDS) Group // Lancet. — 1998. — Vol. 352. — P. 854—865.
62. UK Prospective Diabetes Study (UKPDS) Group // Ibid. — P. 837—853.
63. UK Prospective Diabetes Study (UKPDS) Group // Br. Med. J. — 1998. — Vol. 317. — P. 703—713.
64. Uusitupa M., Laitinen J., Siitonen O. et al. // Diabetes Res. Clin. Pract. — 1993. — Vol. 19. — P. 227—238.
65. Von Korff M., Gruman J., Schaefer J. et al. // Ann. Intern. Med. — 1997. — Vol. 127. — P. 1097—1102.
66. Wysocki T. // J. Pediatr. Psychol. — 1993. — Vol. 18. — P. 443—454.
67. Wysocki T., Greco P., Harris M. A. et al. // Diabetes Care. — 2001. — Vol. 24. — P. 441—446.

Поступила 10.01.03

◆ СОВРЕМЕННЫЕ ПРОБЛЕМЫ МЕДИЦИНСКОЙ НАУКИ

© КОЛЛЕКТИВ АВТОРОВ, 2003

УДК 61:001.8

В. В. Фадеев, В. П. Леонов, О. Ю. Реброва, Г. А. Мельниченко

ДОКАЗАТЕЛЬНАЯ МЕДИЦИНА И ОТЕЧЕСТВЕННАЯ МЕДИЦИНСКАЯ НАУКА

Кафедра эндокринологии (зав. — акад. РАН И. И. Дедов) ММА им. И. М. Сеченова, факультет информатики Томского государственного университета, НИИ неврологии (дир. — акад. РАМН Н. В. Верещагин) РАМН, Москва

В прошедшем XX веке клиническая медицина сделала ошеломляющие успехи: были созданы антибиотики, вакцины, препараты гормонов, гипотензивные и противоопухолевые средства, которые спасли и продлили жизнь миллионам людей. Головокружение от успехов проявилось прогрессирующим увеличением числа специалистов, занимающихся клинической и теоретической медициной, появилось огромное число медицинских журналов, научных работ и публикаций. Базой для этого явилось тысячекратное увеличение финансирования медицины. Так, в 1991 г. в США расходы на медицину составили 752 млрд долларов, или 13,2% от валового национального продукта.

В связи с этим в обществе, тратящем на медицинскую науку колоссальные средства, возник вопрос: "А как именно тратятся эти деньги, кем они тратятся и кому именно эти деньги выделять в дальнейшем?". С другой стороны, у самой медицинской общественности и, что не менее важно, у организаторов здравоохранения возникли вопросы: "Результатам каких именно работ следует доверять? Какие результаты использовать в качестве клинических рекомендаций для практического здравоохранения?". Для ответа на эти вопросы в 80-х годах XX века была разработана концепция доказательной медицины (evidence-based medicine), которую пионеры клинической эпидемиологии Р. и С. Флетчеры [10] определяют как медицинскую практику, ос-

нованную на результатах правильно организованных клинических исследований. В соответствии с концепцией доказательной медицины каждое клиническое решение врача должно базироваться на научных фактах, а "вес" каждого факта становится тем больше, чем строже методика исследования, в ходе которого этот факт получен.

При обсуждении принципов доказательной медицины можно услышать возражение о том, что они ставят под сомнение значение всех тех медицинских значений, которые накапливались тысячелетиями. Однако доказательная медицина подразумевает более совершенную методологию исследования медицинских вмешательств с неочевидной эффективностью, соответствующую современному планированию научного эксперимента. Действительно, эффективность жгута при кровотечениях доказать достаточно просто: после его наложения оно останавливается. Очевидно, что это вмешательство эффективно. Аналогичным образом обстоит дело с заместительной гормональной терапией. Классическая экспериментальная эндокринология XIX века базировалась на удалении эндокринной железы и демонстрации регресса развившихся изменений путем введения в организм экспериментального животного экстракта этой железы. Так же исчерпывающе доказателен эффект инсулина при лечении сахарного диабета типа I и препаратов левотироксина при лечении гипотиреоза. В данном случае теоретические представления исследователя о причинах нарушения функций совпадают с объективными патофизиологическими фактами и поэтому подтверждаются клинической практикой.

Значительно сложнее проблема различных лечебных вмешательств обстоит в тех случаях, когда осуществляется попытка воздействия на патологический процесс, представляющий собой каскад взаимосвязанных биохимических реакций, развертывающихся в самой клетке и в окружающих ее структурах. При этом динамика цепных реакций, сигналы, определяющие их начало и завершение, как правило, остаются неизвестными. В результате представления о патологическом процессе, являющемся целостной динамической системой, сводятся к частностям: фрагментам и звеньям отдельных биохимических реакций, на которые нередко направлены многочисленные фармакологические влияния. В этих случаях реализация теоретических предположений, основанных даже на высокотехнологичных экспериментальных данных, не дает явного клинического эффекта. Так, например, дело обстоит с применением различных классов ангиопротекторов при атеросклерозе, клинические испытания которых в рамках рандомизированных исследований пока не привели к положительному результату. В концепции доказательной медицины теоретические представления о патогенезе (обычно неполные), мнения авторитетных специалистов, традиции, личный опыт, соображения приоритетности не считаются убедительными, научными основаниями для использования того или иного медицинского вмешательства. Ценность теоретических представлений о свойствах лекарственных препаратов бесспорна, вместе с тем их истинные терапевтические возможности можно определить лишь в результате клинических испытаний, выполненных по правилам, принятым международным медицинским сообществом.

В журнале "British Medical Journal" [12] была опубликована статья с шуточным перечислением вариантов "медицинской практики, не основанной на доказательных фактах", в противовес доказательной медицине. Авторы указывают на то, что если медицинская практика не базируется на доказанных фактах, то в ее основе лежит следующее: 1) авторитет врача ("увеличение числа однотипных ошибок с увеличением стажа работы"); 2) страстность ("эмоциональное воздействие на более спокойных коллег и родственников больных"); 3) внешний облик и красноречие ("хороший загар, шелковый галстук, вальяжная поза и красноречие как замена доказанным фактам"); 4) провидение ("когда неизвестно, что делать с больным, вместо обоснованного решения полагаются на волю божью"); 5) чувство неуверенности ("от чувства растерянности и отчаяния решения вовсе не принимаются"); 6) нервозность ("в условиях постоянного страха перед судебным процессом врач назначает чрезмерное обследование и лечение"); 7) самоуверенность ("в основном для хирургов").

По современным стандартам, надежная оценка эффективности методов лечения и профилактики может быть получена только в ходе рандомизированных контролируемых испытаний. Такого рода испытания проводятся не только для фармакологических препаратов, но и для хирургических методов лечения, физиотерапевтических процедур, мероприятий в области организации медицинской помощи, диагностических методов. По окончании исследования сопоставляются частоты наступления

клинически важных исходов — выздоровления, осложнения, смерти, а не суррогатные исходы — изменения физиологических, биохимических, иммунологических и других показателей. Таким образом, оцениваются не столько непосредственные результаты (лабораторные, инструментальные и др.), сколько отдаленные (клинические эффекты, установленные в качестве конечных точек исследования).

Для получения выводов исследования необходимо учитывать неопределенность многих характеристик, а также конечность числа наблюдений. Наиболее приемлемым инструментом в этом случае оказываются методы статистики. Именно эту особенность и подчеркивает одно из определений статистики, которое было дано американским математиком А. Вальдом: "статистика — это совокупность методов, которые дают нам возможность принимать оптимальные решения в условиях неопределенности".

Почему отечественная медицина пока не пошла по пути доказательной медицины, а работы отечественных авторов стоят особняком и казуистически редко цитируются за рубежом? Почему отечественные медицинские журналы имеют столь низкий рейтинг на фоне зарубежных? Причин, наверное, много.

Во-первых, отечественная медицина не испытывала на себе столь мощные экономические стимулы и никогда не финансировалась в тех объемах, о которых сказано выше. Вторым безусловным фактором является длительная изоляция советской медицинской науки от мировой, причем отчасти эта изоляция имела демонстративный характер, а сами работы отечественных авторов противопоставлялись зарубежным. До сих пор список литературы к статье или диссертации многими воспринимается не как список процитированной литературы, а как некое "самостоятельное произведение", в котором определенная доля источников должна принадлежать отечественным авторам. Парадоксально, но наряду с упреками в адрес списка литературы в ряде случаев можно не услышать никаких замечаний в адрес самой работы.

Тот факт, что все без исключения отечественные медицинские журналы издаются на русском языке, не оправдывает факт низкого рейтинга этих журналов в научном мире. В этом плане хорошим примером являются издаваемые на немецком языке журналы "Deutsche Medizinische Wochenschrift" или "Acta Medica Austriaca", которые имеют высокий международный рейтинг, а порой традиционно более полно освещают некоторые проблемы.

Следствием указанных факторов явилось серьезное несоответствие методологической базы российской медицинской науки современным международным стандартам получения научно обоснованных, доказательных результатов исследований. По современным представлениям, таковыми являются результаты исследований, отвечающих следующим требованиям: 1) соответствие способа организации исследования поставленным задачам; 2) минимизация систематических ошибок, возникающих при формировании несбалансированных групп наблюдений, вследствие неправильных и нестандартизированных измерений, из-за вмешивающихся факторов; 3) минимизация случайных ошибок путем корректного анализа данных.

Наличие в исследовании систематических ошибок и несоответствие способа организации исследования поставленным задачам делает исследование бессмысленным. Никакой, даже самый совершенный анализ данных не может компенсировать дефекты в организации исследования. В то же время правильно организованное исследование, в рамках которого анализ данных проведен с ошибками, также сводит на нет все усилия исследователей. Таким образом, первичной проблемой является правильная организация исследования, а вторичной — корректный анализ данных. В настоящее время считается, что неправильно проведенное (в смысле как его организации, так и анализа данных) исследование является неэтичным по крайней мере по следующим основаниям: пациенты в ходе исследования подвергаются неоправданному риску; происходит неэффективное использование ресурсов (финансов, времени исследователей), которые могли бы быть потрачены на преодоление более важных проблем; после публикации неверных результатов дальнейшие исследования направляются в неправильное русло; применение неверных результатов исследования в медицинской практике способно нанести вред пациентам.

Безусловно, наибольшие трудности (как организационные, так и финансовые) представляет собой правильная организация исследования. Этими проблемами занимается наука "клиническая эпидемиология" — методологическая база доказательной медицины, основы которой были заложены канадскими учеными D. Sackett, B. Haynes, G. Guyatt, P. Tugwell в 1980-е годы. На-

личие хорошей литературы на русском языке (прежде всего [10]) позволяет нам подробно не останавливаться в данной публикации на весьма сложных вопросах организации доказательного медицинского исследования.

Как уже было сказано выше, не меньшее значение для получения научно обоснованных результатов имеет корректность статистического анализа данных. Клиническая статистика, как и любая другая медицинская специальность, является отдельной дисциплиной. Конечно, врач не в состоянии в полной мере овладеть всеми медицинскими специальностями, которые необходимо привлекать в конкретном исследовании, например лабораторной диагностикой или рентгенорадиологией. Как нам представляется, оптимальным вариантом является участие в исследовании профессионального статистика. Однако поскольку число таких специалистов в России крайне мало, исследователям необходимо овладеть хотя бы основами статистического анализа данных.

Источники научно обоснованной информации

"Клиническая эпидемиология — методическая основа доказательной медицины — повышает эффективность восприятия информации, поскольку врач может, исходя из фундаментальных принципов, быстро разобраться, какие статьи или источники информации заслуживают доверия. Это обеспечивает достойную альтернативу тем признакам, по которым обычно судят о правоте собеседника в процессе общения с коллегами — убежденности, красноречию, профессиональному стажу или глубине специализации" [10].

В 1972 г. британский эпидемиолог А. Кокран обратил внимание на то, что "общество пребывает в неведении относительно истинной эффективности лечебных вмешательств. Очень стыдно, что медики до сих пор не создали системы аналитического обобщения всех актуальных рандомизированных клинических испытаний по всем дисциплинам и специальностям с периодическим обновлением обзоров". Он предложил создавать научные медицинские обзоры на основе систематизированного сбора и анализа фактов, а затем регулярно пополнять их новыми данными.

В октябре 1992 г. был открыт первый Кокрановский центр в Оксфорде. В октябре 1993 г. состоялся первый ежегодный Кокрановский colloquium, где 73 представителя из 9 стран стали соучредителями Кокрановского сотрудничества, основной задачей которого является создание, обновление и распространение систематических обзоров результатов медицинских вмешательств, которые должны облегчить принятие решений в различных областях медицины. Что же в настоящее время понимают под систематическим обзором?

Совершенно очевидно, что поток информации о результатах медицинских исследований и публикаций огромен. Даже если заниматься каким-нибудь узким вопросом в какой-либо узкой области медицины, то одному человеку переработать всю информацию, особенно вечером, после бурно проведенного рабочего дня, невозможно. Например, если заниматься относительно узкой областью медицины — эндокринологией, далее еще более узкой областью эндокринологии — тиреологией, далее еще больше сузиться и заниматься исключительно проблемой узлового зоба, то на ключевые слова "узловые образования щитовидной железы" ("thyroid nodules") база данных MEDLINE выдает список из нескольких десятков тысяч статей. Однако и это еще не вся проблема. Результаты этих работ, как правило, противоречат друг другу. Ведь при проведении исследований авторы руководствовались разными принципами набора групп, включали в них разное количество пациентов, которые различались по полу, возрасту, национальности, отношению к этому исследованию и т. д. Разные авторы "хотели доказать" и "доказывали" разные факты. Лишь небольшая часть исследований является рандомизированными и контролируруемыми. Этот список причин гетерогенности исследований можно продолжать. И все это падает на голову человека, который формирует на основе этого материала свое субъективное мнение (причем зачастую смещенное, так как в его распоряжении могут оказаться лишь некоторые работы, в частности работы с положительными результатами из-за *тенденции исследователей не публиковать отрицательные результаты*) и проводит свое исследование, которое, как правило, это мнение подтверждает.

Систематические обзоры призваны как-то разрешить эту проблему. Такие обзоры — это научные исследования с заранее спланированными методами, где объектом изучения служат результаты ряда оригинальных исследований. Они синтезируют результаты этих исследований, используя подходы, уменьшаю-

щие возможность систематических и случайных ошибок. Эти подходы включают в себя максимально глубокий поиск публикаций по определенному вопросу и использование точных, воспроизводимых критериев отбора статей для обзора. Далее производится оценка структуры и особенностей оригинальных исследований, обобщение данных и интерпретация полученных результатов. Если результаты оригинальных исследований рассмотрены, но статистически не объединены, обзор называется качественным систематическим обзором. В количественном систематическом обзоре, иначе называемом *метаанализом*, для объединения результатов двух или более исследований используются статистические методы [11]. Объединяют не все подряд исследования, а лишь качественно выполненные, предпочтительно рандомизированные и сопоставимые по материалу. Исследования, которые базируются лишь на "собственном видении проблемы", в которых отсутствует контрольная группа, а статистический анализ представляет собой лишь указание в скобках на то, что $p < 0,05$, объединить невозможно (см. ниже).

Систематические обзоры лишь отдаленно напоминают те обзоры, которые мы привыкли видеть в журналах и диссертационных работах. Обычные обзоры рассматривают широкий круг проблем. Например, обычный обзор может называться "Узловой зоб", тогда как систематический обзор будет называться "Эффективность супрессивной терапии левотироксина при узловом коллоидном зобе небольшого размера". В систематическом обзоре все силы брошены для поиска ответа на достаточно узкий клинический вопрос, обычно об эффективности клинического вмешательства, а не просто на изложение клинической проблемы, напоминающее главу из учебника.

Деятельностью по подготовке систематических обзоров занимается всемирное Кокрановское сотрудничество врачей. В настоящее время подготовлено 1000 (по состоянию на март 2001 г.) обзоров по всем областям медицины. Число обзоров быстро растет — примерно на 50% в год. Кокрановское сотрудничество формирует также реферативную базу данных, включающую в себя рефераты публикаций о контролируемых и рандомизированных клинических испытаниях, отвечающих современным стандартам качества их проведения, результаты которых, следовательно, являются научно обоснованными, доказательными. Кокрановская рабочая группа по проблемам метаболических и эндокринных заболеваний к началу 2001 г. выполнила 4 систематических обзора и зарегистрировала 9 протоколов обзоров, подготавливаемых в настоящее время и посвященных сахарному диабету. Рабочей группой по проблемам беременности выполнен обзор "Йодная профилактика у беременных в йоддефицитных регионах" [13].

Информированность отечественных исследователей о результатах научно обоснованных клинических испытаний

В последние годы получение научной информации существенно упростилось за счет доступности ресурсов Интернета. Наличие бесплатного (для пользователей Интернета) доступа к реферативной базе данных MEDLINE национальной медицинской библиотеки США и полнотекстовым электронным версиям ряда периодических изданий позволяет ожидать повышения информированности российских исследователей о проводящихся за рубежом клинических испытаниях и их публикуемых результатах.

Казалось бы, такая осведомленность должна проявляться в цитировании этих результатов в отечественных публикациях. Изучение этого вопроса показало обратное. В недавней работе [9] представлен анализ 44 оригинальных статей, опубликованных в крупнейших отечественных медицинских журналах. Из заголовка, текста и(или) списка ключевых слов статей были выделены ключевые термины для проведения поиска в базе контролируемых исследований Кокрановского сотрудничества, содержание которой практически полностью отражено в базе MEDLINE. После этого сопоставляли количество работ, процитированных в оригинальных статьях отечественными авторами, с количеством контролируемых клинических исследований, которые по сочетанию терминов удалось найти в базе данных. В итоге оказалось, что в 44 проанализированных работах теоретически могло быть процитировано 706 релевантных источников. В действительности же было процитировано всего 7 (т. е. около 1%) рандомизированных исследований.

В значительной мере анализ осложнялся тем, что в отечественных работах часто (в 7 из 44) изучали эффективность препаратов при состояниях, отличающихся от состояний, являющихся показаниями к применению данных препаратов в зарубежных исследованиях (например, "infrared laser + diabetic polyneu-

ropathy", "lung cancer + plasmapheresis", "sulodexide + diabetic neuropathy"). Ряд статей (5 из 44) используют термины, практически не применяемые в зарубежной литературе (например, "pseudoallergy"). Пять статей были посвящены исследованию отечественных препаратов, которые пока не изучаются за рубежом (Salben, Likopid, Prospedin, Urosan, Kerlon).

В ряде статей, несмотря на значительное количество (от 21 до 134) соответствующих зарубежных публикаций, они не цитируются или цитируются недостаточно (10 из 44). Для большинства же проанализированных статей (20 из 44) число соответствующих источников невелико — менее 10.

Таким образом, наличие недостаточное цитирование имеющихся публикаций. Такое положение, возможно, связано с тем, что авторы могут знать, но не цитировать аналогичные зарубежные работы, преувеличивая тем самым новизну и актуальность своих исследований. Не менее вероятно, что отечественные врачи просто не знают о существующих зарубежных публикациях.

Использование статистических методов в отечественных медицинских исследованиях

На сегодняшний день более или менее очевиден тот факт, что доказательство каких-либо гипотез в медицине, как правило, требует применения статистических методов. Почему этот факт очевиден именно "более или менее", мы покажем дальше. Здесь хочется обратить внимание на то, что еще в конце XIX — начале XX века российская медицинская наука активно осваивала статистический инструментарий [5]. Однако в период 40-х—50-х годов гонения на генетику привели к изгнанию из биологии и медицины и методов статистики [5]. В 1940 г. в "Сборнике докладов академии наук СССР" академик Т. Д. Лысенко пишет: "Я не чувствую себя достаточно компетентным, чтобы разбираться в этой системе математических доказательств. К тому же меня, как биолога, сейчас не интересует вопрос о том, хорошим или плохим математиком был Мендель... Нас, биологов, не интересуют математические выкладки, подтверждающие практически бесполезные статистические формулы менделистов" [6]. Вслед за статьей Т. Д. Лысенко следует статья Э. Колымана [3], цитаты из которой, на наш взгляд, продолжают до настоящего времени определять отношение к статистике авторов многих отечественных работ: "Мы, биологи, не желаем подчиняться слепой случайности (хотя бы математически и допустимой) и утверждаем, что биологические закономерности нельзя подменять математическими формулами и кривыми. Теория вероятностей и статистический метод исследования являются лишь вспомогательными орудиями в конкретной науке (например, в политической экономии, в физике, в биологии). В зависимости от того, *какая конкретная теория контролирует ее применение, статистика будет давать результаты, правильно или неправильно отражающие материальную действительность*". Если перефразировать Э. Колымана, получится знакомая многим шутка: "Статистика должна служить человеку, а не человек статистике". Отсюда формируется стандартный алгоритм исследования. Аспирант или докторант получает тему, начинает над ней активно работать, приходит к неким выводам. Далее возникает вопрос о подтверждении этих выводов, т. е. исследователь изыскивает метод, который подтвердит уже сделанные им выводы, а не наоборот: использует заранее запланированный статистический метод, который подтвердит или отвергнет заранее сформулированную гипотезу. Более того, сама тема, которая поручается аспиранту, очень часто уже подразумевает некий вывод, который формулируется еще до выполнения работы. В связи с этим вспоминается работа, которая, после рассуждений об актуальности темы, начиналась словами: "Поскольку мы уже тридцать лет оперируем, мы имеем право на собственное мнение". Парадокс, но права на собственное мнение никто не отменял. Зачем же эта фраза? Авторы чего-то боялись, пуская в ход даты окончания медицинского института вместо строгой научной аргументации своего мнения? Таким образом, традиции последователей академика Лысенко ("конкретная теория контролирует применение статистики") продолжают жить.

Большинство используемых в современной медицине статистических методов было разработано еще в начале XX века. Так, всеми любимый *t*-критерий Стьюдента был разработан Вильямом Госсетом в 1908 г. Широко используемые в настоящее время непараметрические методы были разработаны позднее — в 40-х—50-х годах. Во-вторых, следует заметить, что большинство наиболее используемых ныне статистических методов было подробно описано в отечественной литературе прошлых лет. Более того, в 70-х годах издавалось значительно больше книг по использованию статистики в медицине и планированию экспе-

римента. Так, аргументация преимуществ непараметрических методов, использующихся в отечественных работах в минимальном объеме, была подробно и доступно описана в книге "Применение непараметрических критериев статистики в медико-биологических исследованиях", которая вышла в 1973 г. [2]. В дореволюционной России выходило большое количество руководств по биологической статистике, а знание этой дисциплины активно пропагандировалось [5]. Так, в своем учебнике по основам военно-полевой хирургии Н. И. Пирогов пишет: "Я принадлежу к ревностным сторонникам рациональной статистики и верю, что приложение ее к военной хирургии есть несомненный прогресс" [7].

По мнению А. Е. Платонова [8], ученый-медик в настоящее время может не знать ни одной формулы, но должен четко понимать следующее:

- как формулируется тот вопрос, на который он хочет ответить с помощью статистического анализа;
- какие статистические процедуры необходимы для ответа на данный вопрос;
- как следует интерпретировать результаты.

С одной стороны, действительно, возможности современных статистических пакетов программ во многом избавляют исследователей от необходимости знания статистических формул. Однако даже для того, чтобы сформулировать правильно вопрос, на который ученый-медик хочет ответить с помощью статистического анализа, ему необходимо понимание как статистической терминологии, так и методов поиска ответа на эти вопросы. Можно ли, не зная ни одной формулы по статистике, выбрать адекватную статистическую процедуру и затем столь же адекватно интерпретировать результат ее применения? Аналогичной является ситуация, когда человеку, не знающему анатомии человека, предложили бы провести хирургическую операцию и по ее окончании дать прогноз состояния здоровья пациента.

Как же обстоит дело с описанием использования статистических методов в отечественной биологии и медицине? В статье [4] приводятся результаты анализа 1538 отечественных оригинальных статей и 259 диссертационных работ по биологии и медицине, которые были опубликованы с 1990 по 1997 г. Почти в каждой четвертой из работ вообще отсутствовало упоминание о статистических методах и критериях. Тем не менее большинство таких работ содержат в тексте и таблицах упоминание о том, что $p < 0,05$ или $p < 0,01$. Среди оставшихся работ названия компьютерных программ, с помощью которых проводился анализ, указаны только в 3% случаев. В большинстве работ, в которых используется статистическая обработка данных, этим упоминанием, как правило, и ограничиваются. В 45% работ такое упоминание включает использование клише типа "результаты обрабатывались статистически", "обработку данных проводили статистическими методами", "результаты исследований подвергли обработке общепринятыми методами вариационной статистики" и т. д. Если провести аналогию на более понятном клиникусту языке, эти формулировки напоминают фразу "больных лечили стандартными (общепринятыми) методами" без упоминания самого стандарта. Еще в 25% проанализированных работ описание применявшихся статистических методов безграмотно или противоречиво ("вероятность случайности различий соответствует достоверности", "различия считались достоверным при $p < 0,05$, т. е. в тех случаях, когда вероятность различия составляла больше 95%"). В 85% работ при описании выборок авторы использовали выражение " $M \pm m$ ", при этом в большинстве случаев без объяснения того, что обозначают эти буквы: если M — это скорее всего среднее арифметическое, то что такое m (среднеквадратичное отклонение, стандартная ошибка среднего и т. д.) — неясно. После уточнения у автора в большинстве случаев оказывается, что это "что-то $\pm$ что-то" является "среднее арифметическое $\pm$ его стандартная ошибка". Пришедшее в биологическую статистику из физики широкое использование стандартной ошибки среднего в большинстве случаев не оправдано, поскольку не позволяет читателю сделать вывод о реальном разбросе в выборке, т. е. о ее однородности.

Результаты изучения частоты использования различных статистических критериев оказались еще более мрачными: в 15% случаев статистический анализ не проводили, а в 53% случаев использованный статистический критерий не указан. В тех случаях, когда было сделано и то и другое, в 59% работ был использован критерий Стьюдента. Корреляционный анализ проводили в 9% работ, критерий "Хи-квадрат" использован в 9% работ, 4% пришлось на дисперсионный (ANOVA) и 4% — на регрессионный анализ. Другие методы использовали лишь в 9% случаев.

С применением *t*-критерия Стьюдента складывается вполне обычная ситуация, т. е. та, которая описывается в большинстве современных руководств по медицинской статистике в разделе "Типичные ошибки". Большое количество примеров некорректного применения критерия Стьюдента приведено и в специальном разделе "Кунсткамера" электронного журнала "Биометрика" (<http://www.biometrika.tomsk.ru/kk/ndex0.htm>). В недавно переведенной на русский язык и вышедшей в издательстве "Практика" книге С. Гланца "Медико-биологическая статистика" [1] автор прямо в предисловии пишет: "... Результаты многих биологических и медицинских работ основаны на неправильном использовании статистических методов и способны только ввести в заблуждение. Большинство ошибок связано с неправомерным использованием критерия Стьюдента".

Как уже указывалось, в 7% проанализированных работ использовался корреляционный анализ. В ситуации, когда статистический анализ начинается после того, как работа уже выполнена, исследователю бывает трудно удержаться от того, чтобы не попытаться "прокоррелировать" все и вся — "а вдруг обнаружится какая-то взаимосвязь?" Такая постановка вопроса при всей ее заманчивости аналогична описанной выше, когда автор, проводя многочисленные попарные сравнения, пытается ответить на вопрос: "А нет ли между какими-нибудь группами хоть какого-нибудь различия или связи?" При критическом уровне статистической значимости $p = 0,05$ весьма вероятно, что для одного из нескольких десятков сравнений достигнутый уровень значимости окажется менее 5%. Наличие корреляции еще не означает, что между этими явлениями присутствует причинно-следственная связь, и, более того, отсутствие корреляции не означает, что эта связь отсутствует.

В каждом втором случае использования корреляционного анализа авторы не указывают, какой именно коэффициент корреляции был ими использован — Пирсона, Спирмена или какой-либо другой [4]. Кроме того, достаточно типичным является использование авторами статей оборотов типа "выявлена достоверная корреляция" без приведения конкретных результатов корреляционного анализа, т. е. термин "корреляция" очень часто используется не как определенное статистическое понятие, а как характеристика взаимосвязи явлений. Довольно бессмысленно и само использование словосочетания "достоверная корреляция" в силу противоречивости его элементов. Термин "корреляция" предполагает наличие статистической связи, тогда как "достоверной" может быть только функциональная связь, но никак не статистическая. Именно поэтому в статистике используется более корректное выражение "статистически значимая связь" (в том числе и связь корреляционная).

Хотелось бы немного остановиться еще на одной проблеме, которая охватывает не только статистическую обработку данных, но и планирование эксперимента. Как показывает практика, большинство исследователей уверены, что наиболее достоверные данные удастся получить максимальным увеличением объема выборки, а все неудачи и "недостоверные" различия между выборками склонны списывать на недостаточный объем выборки. Лишь в единичных работах приводится расчет необходимого объема выборки, при котором с определенным уровнем значимости (например, 0,05 или 0,01) можно будет отвергнуть гипотезу (нулевую гипотезу) об отсутствии различий определенной величины между выборками. Объем выборки может быть рассчитан, исходя из выбранного уровня значимости (например, 0,05), клинически значимой разницы между исследуемыми признаками и их стандартного отклонения, которые можно оценить в ходе пилотного исследования или при анализе данных литературы. Например, если клинически значимым считать уменьшение объема щитовидной железы на фоне лечения зоба на 5 мл при том, что стандартное отклонение объема щитовидной железы составляет, к примеру, около 3 мл, при желаемом уровне значимости в 0,001 в каждую из двух сравниваемых групп должно войти 15–20 человек. Если, например, в таком исследовании в каждую из групп включить тысячу человек, то вероятность обнаружить статистически значимые различия между группами значительно увеличивается. Однако при этом многократно увеличивается вероятность обнаружить клинически незначимые различия. Другими словами, при столь больших выборках разница между двумя группами может оказаться статистически значимой при том, что объем щитовидной железы на фоне лечения изменится на десятки доли миллилитра, т. е. указанная динамика не будет иметь ни малейшего клинического значения и скорее всего будет определяться погрешностью метода, т. е. в данном примере погрешностью ультразвуковой волнометрии щитовидной железы.

Возвращаясь к началу этого раздела статьи, заметим, что если целью работы является статистическое подтверждение сделанных заранее выводов, то использование приема максимального возможного увеличения объема выборки может привести к тому, что искомые статистические различия будут найдены. Если же исследователь ставит перед собой вопрос о клинически значимых различиях или динамике, а не о различии или динамике вообще, при этом он как истинный ученый способен абсолютно одинаково отнестись как к подтверждению, так и к отвержению сформулированной им заранее гипотезы, то ему следует рассчитать объем выборки и привести эти расчеты в своей статье.

Мы указали лишь на наиболее распространенные проблемы статистической обработки данных медико-биологических исследований. В заключение этого раздела хотелось бы привести еще одну цитату из книги С. Гланца "Медико-биологическая статистика" [1]. С. Гланц работает внештатным редактором "Journal of the American College of Cardiology", и его работа состоит в выявлении статистических ошибок в направляемых в этот журнал статьях. В настоящее время в большинстве зарубежных журналов с высоким рейтингом статьи проходят рецензирование специалистом по статистике, чего нельзя сказать об отечественных биомедицинских журналах и диссертациях. Так вот, С. Гланц ставит в заслугу своей многолетней работе по пропаганде доказательной медицины то, что "доля статей, содержащих ошибки, как и раньше, составляет около половины, но теперь уже половина предлагаемых к публикации, а не опубликованных работ".

Влияние концепции доказательной медицины на медицинскую практику, науку, образование за рубежом и в России

За рубежом концепция доказательной медицины получила распространение не только среди исследователей в области клинической медицины, но и среди практических врачей. По опросу 1996 г. врачи общей практики в Великобритании до 80% клинических решений принимают в соответствии с принципами доказательной медицины. Профессиональными врачебными ассоциациями и группами экспертов разрабатываются научно обоснованные клинические рекомендации по определенным проблемам. Фармацевтические компании используют результаты систематических обзоров в качестве аргументов для включения своих препаратов в национальный формуляр.

Достижения доказательной медицины ощутимо влияют на политику в области научных исследований и образования. Многие ведущие медицинские университеты ввели курс клинической эпидемиологии (науки, являющейся основой концепции доказательной медицины) в обязательную программу в качестве одной из фундаментальных дисциплин. Крупнейшие международные медицинские научные журналы ужесточают требования к публикациям. Ряд изданий, основанных на концепции доказательной медицины ("ACP Journal Club", "Evidence-Based Medicine"), публикуют структурированные рефераты лучших, отобранных по специальным критериям статей из наиболее цитруемых медицинских журналов.

В России пионером пропаганды идей доказательной медицины стал издававшийся с 1996 г. "Международный журнал медицинской практики". Было организовано Российское отделение Кокрановского сотрудничества (<http://www.cochrane.ru/>), содействующее российским ученым в получении информации по проблемам доказательной медицины и в участии в подготовке систематических обзоров.

Несмотря на это, ситуация с внедрением концепции доказательной медицины в науку и практику российской медицины меняется очень медленно. Требования редакций российских медицинских журналов не соответствуют международным. Публикуется значительное количество малодоказательных статей, что связано с неудовлетворительной организацией научных исследований. Очевидно, что в сложившейся обстановке невозможно сразу поднять планку требований до необходимого уровня. Если это сделать быстро и радикально, журналам будет практически нечего публиковать, кроме обзоров и лекций. В этом плане журналы должны выработать последовательную политику.

В России по-прежнему практически не организуются рандомизированные клинические исследования (исключения составляют работы, запланированные зарубежными исследователями, с включением российских учреждений в мультицентровые международные исследования). Все это резко снижает ценность получаемых результатов. Тенденция к широкому распространению в последние годы в России методов лечения и диагностики с научно не доказанной эффективностью, а в худшем

случае — и вредных для здоровья пациентов свидетельствует о низком иммунитете российских врачей к рекламируемым сомнительным методам диагностики и лечения.

Выводы

Приведенные выше данные свидетельствуют о *глубоком кризисе методологии отечественной биомедицинской науки*. При этом здесь мы имеем в виду не плачевное состояние здравоохранения и низкий уровень медицинской помощи населению, а серьезную *методологическую отсталость* выполняющихся в нашей стране исследований, которые в подавляющем большинстве случаев не соответствуют международным стандартам. S. Tilghast, американский врач-эксперт, участвовавший в подготовке программы реформирования здравоохранения в России, заметил, что *"развитие российской медицины будет зависеть от западных знаний и технологий до тех пор, пока современные методы критической оценки информации не изменят культуру медицинских исследований, образования и научных публикаций"*.

Продолжая эту мысль, следует подчеркнуть, что, на наш взгляд, вопреки бытующим представлениям, для отечественной медицинской науки значительно большей проблемой является *не дефицит финансирования, а тот методологический кризис, о котором шла речь в этой статье*. В масштабах страны колоссальные средства тратятся на организацию заведомо некорректно организованных исследований, на финансирование проектов, подразумевающих использование лечебных воздействий с сомнительной или недоказанной эффективностью. Наряду с этим даже простое исследование, проведенное на относительно небольшой группе пациентов, но построенное на принципах доказательной медицины, окажется немногим дороже.

Именно этим, на наш взгляд, и определяется отношение к использованию стандартов диагностики и лечения в широкой клинической практике. Стандарты диагностики и лечения, которые существуют в большинстве высокоразвитых стран, базируются исключительно на исследованиях, проведенных на принципах доказательной медицины. Отсутствие же представлений об этих принципах приводит к непониманию необходимости использования стандартов диагностики и лечения, на основании которых в свою очередь должно базироваться планирование здравоохранения на государственном уровне. Это наиболее яркий пример того, как кризис в науке обуславливает экономический кризис целой отрасли.

Глубокий кризис охватывает большинство описанных звеньев системы здравоохранения и медицинской науки, начиная от обучения в медицинских вузах, заканчивая сертификацией врачей и планированием здравоохранения. Первоочередными мерами по выходу из сложившегося кризиса, как нам кажется, могли бы быть следующие.

1. Введение независимого рецензирования и экспертизы (в том числе статистического) научной обоснованности выводов биомедицинских статей и диссертационных работ путем введения специальных экспертов в состав редакций журналов, диссертационных советов и экспертных советов ВАК РФ.

2. Создание специализированных журналов по биометрике и медицинской статистике, а также редакционных разделов в основных журналах, где будут печататься аналогичные, но более специализированные материалы.

3. Преподавание основ доказательной медицины и статистики во всех медицинских и биологических вузах и включение экзаменов по этим предметам в обязательный кандидатский минимум.

4. Организация и расширение полномочий лабораторий биостатистики и информатики в биомедицинских НИИ и вузах.

5. Активная пропаганда принципов доказательной медицины на всех этапах медицинского образования.

ЛИТЕРАТУРА

1. Гланц С. Медико-биологическая статистика: Пер. с англ. — М.
2. Гублер Е. В., Генкин А. А. Применение непараметрических критериев статистики в медико-биологических исследованиях. — Л., 1973.
3. Кольман Э. // Докл. АН СССР. — 1940. — Т. 28, вып. 1. — С. 836—840.
4. Леонов В. П. // Междунард. журн. мед. практики. — 1998. — № 4. — С. 7—12.
5. Леонов В. П. // Там же. — 1999. — № 4. — С. 7—19.
6. Лысенко Т. Д. // Докл. АН СССР. — 1940. — Т. 28, вып. 1. — С. 834—835.
7. Пирогов Н. И. Начала общей военно-полевой хирургии. — М.-Л., 1941. — Ч. 1. — С. 1—2.
8. Платонов А. Е. Статистический анализ в медицине и биологии. — М., 2000.
9. Реброва О. Ю. // Междунард. журн. мед. практики. — 2000. — № 8. — С. 5—8.
10. Флетчер Р., Флетчер С., Вагнер Э. Клиническая эпидемиология. Основы доказательной медицины: Пер. с англ. — М., 1998.
11. Cook D. J., Mulrow C. D., Haynes R. B. // Ann. Intern. Med. — 1997. — Vol. 126. — P. 376—380.
12. Isaacs D., Fitzgerald D. // Br. Med. J. — 1998. — Vol. 316. — P. 361—365.
13. Mahomed K., Gülmezoglu A. M. // The Cochrane Library. — Oxford, 2000. — Issue 4.

Поступила 29.10.2000