

Возможности и ограничения метаанализа

М. Ферстрате¹

Центр молекулярной и сосудистой биологии
Университет г. Лювена, Бельгия

Ключевые слова: метаанализ, консолидация данных, клинические исследования, кумулятивный метаанализ.

Метаанализ — ценный метод консолидации данных ряда различных исследований, которые сами по себе в отдельности недостаточны или неубедительны. Есть четкие правила поиска и сбора информации, которые являются принципиально важными для обеспечения достоверности метаанализа. Многие попытки так называемого метаанализа безуспешны по причине отсутствия протокола или несоответствия ему. Неполные данные об исследованиях влияют на достоверность метаанализа. К примеру, утаивание информации — неслучайное явление, в основном оно касается исследований с отрицательными результатами. Во избежание таких часто встречающихся ошибок, связанных с преимущественной публикацией только положительных результатов исследований, все клинические наблюдения должны быть внесены в международную он-лайн базу данных.

Метаанализ можно определить как объединение результатов нескольких рандомизированных контролируемых исследований с аналогичным дизайном для получения единой оценки эффективности лечения, которая является более точной, нежели оценка результатов исследований по отдельности. Этот обзор направлен на борьбу с субъективными и часто содержащими систематические ошибки традиционными повествовательными обзорами, которые в большинстве случаев не упоминают источник информации и не отвечают методологическим стандартам качества исследований (1, 2).

Изначально метаанализ был встречен неоднозначно, так как проведенный таким образом анализ небольших исследований противоречил результатам одного большого рандомизированного исследования по аналогичной теме. Хорошо известными примерами являются: изучение влияния нитратов (3, 4) и магнезии (5, 6) на смертность при остром инфаркте миокарда, общая внутригоспитальная оценка смертности среди пожилых лиц (7,

8) и изучение влияния аспирина на риск развития преэклампсии (9, 10).

Очевидно, что решающим моментом является качество исследований, входящих в метаанализ: если исходный материал неудачный, то это скажется на метаанализе (при использовании не Достоверных исходных данных будет получен не Достоверный результат). Несколько экспертов проанализировали многочисленные систематические ошибки, угрожающие достоверности метаанализа (11, 12). Ошибки могут быть связаны с тем, что входящие в метаанализ исследования отличаются по характеристикам пациентов (систематические ошибки отбора), лечению (систематические ошибки исполнения), оценке результатов (систематические ошибки определения исхода) и исключению рандомизированных пациентов (систематические ошибки, связанные с отсевом).

На достоверность описательных обзоров и метаанализа часто влияет неполнота сообщаемых результатов исследований. Это проявляется тогда, когда лишь около половины сообщений, представленных на конференциях, в дальнейшем публикуются в полном объеме (13). Из всех проектов исследований, одобренных этическими комитетами четырех ведущих медицинских школ, только 49-67 % в полном объеме публикуются в медицинских журналах (14, 15). Результаты 20 % исследований, проводимых под эгидой Национального института здоровья, остаются неопубликованными в течение нескольких лет после их завершения (14, 16). Такое утаивание информации, к несчастью, неслучайное явление — оно касается, в основном, исследований с негативными последствиями. Таким образом, обзор опубликованных данных может давать ложную информацию об эффективности лечения, или в нем может быть упущен важный побочный эффект. Во избежание ошибок, связанных с публикациями (неопубликованные данные или повторная публикация данных), было предложено вносить все исследования в международную (национальную) он-лайн базу данных (17). Кроме Кокрановского регистра контролируемых испытаний существует Pub Med Central — электронный архив биомедицинских исследований со свободным доступом, введенный в действие Национальным институтом здоровья США (18).

Метаанализ — это гораздо больше, чем объединение данных отдельных исследований

Метаанализ не может обеспечить такое же ка-

¹ Center for Molecular and Vascular Biology
Herestraat 49, B-3000 Leuven, Belgium.
Phone: 32-16-345775
Fax: 32-16-345990
e-mail: marc.verstraete@med.kuleuven.ac.be
Статья получена 28 февраля 2004 г.
Принята в печать 12 апреля 2004 г.

чество информации об эффективности воздействия, как единое большое рандомизированное исследование, если будут использованы менее строгие стандарты.

Действующие в настоящее время стандарты подчеркивают важность некоторых качественных показателей, например, таких как разработка протокола перед началом метаанализа, включение только достоверно рандомизированных исследований и сбор полной информации о результатах у всех рандомизированных пациентов (11, 19, 20). Нередко метаанализ оказывался безуспешным по причине отсутствия протокола или несоответствия ему. Другим важным недостатком является то, что метаанализ проводился после того, как становились известны окончательные результаты отдельных исследований. Во избежание таких ошибок во время отбора рандомизированных исследований для включения в метаанализ информация о результатах исследования должна быть скрыта.

Таким образом, процесс поиска и отбора данных является решающим и предопределяет достоверность метаанализа. Необходимо строго придерживаться следующих критериев, предложенных Chalmers и соавт. (19):

Правильная двойная слепая рандомизация

Главные процедуры играют важную роль при любой схеме исследования, так как они позволяют свести к минимуму вероятность систематических ошибок при назначении лечения.

Полная доступность данных в соответствии с принципами анализа всех включенных пациентов (intention to treat)

Полный учет каждого «выбывшего до оценки результатов» пациента и любой другой недостающей информации (такой как процент пропущенных визитов, количество и причины случаев прекращения лечения и т.п.).

Процедуры для обеспечения достоверности результатов

Для расчета размера выборки метаанализа особых правил нет.

Одним из главных предназначений метаанализа является обеспечение достоверной оценки эффекта при использовании типичных методов лечения, если объем собственно рандомизированных клинических исследований недостаточен. Однако, как можно определить, достаточен ли объем самого метаанализа? Как будет оцениваться статистическая достоверность данных в пределах метаанализа? Есть основания предполагать, что размер выборки должен быть, по крайней мере, настолько же большим, как в одном, хорошо спланированном и оптимально мощном контролируемом рандомизированном исследовании. Учитывая неоднородность схемы исследований и другие возможные систематические ошибки отдельных исследований,

включенных в метаанализ, может потребоваться большее количество пациентов, нежели для одного рандомизированного исследования. Pogue и Yusuf (20) предложили формулу для определения оптимального объема информации (ООИ), с помощью которой можно рассчитать в первом приближении минимальный необходимый размер выборки.

На практике расчет объема анализа редко выполняется проспективно, и это объясняет, почему результаты метаанализов и соответствующих больших исследований не всегда согласуются друг с другом (21).

Риск метаанализов, основанных на малых исследованиях

Исследования, построенные на малом количестве их проведения и с номинально значимой величиной p , могут переоценивать степень эффективности лечения, поскольку для получения статистически достоверной значимости результатов при малом размере выборки эффект лечения должен быть впечатляющим. Эта систематическая ошибка усугубляется еще и тем, что исследования с негативным результатом часто не публикуются. Следовательно, метаанализ, включающий, в основном, малые исследования, скорее всего, переоценивает эффект лечения. Информация о более крупных исследованиях доходит до широких кругов международных исследователей, и вероятность их публикаций гораздо больше даже в случае негативных результатов. Таким образом, метаанализ больших исследований менее подвержен систематическим ошибкам, связанным с публикациями (20).

Существует и другая причина того, что данные, полученные при метаанализе малых исследований, даже при убедительных показателях различия (достоверности), должны интерпретироваться с осторожностью (20). Даже в исследованиях, выполненных одинаковым образом в идентичных популяциях, присутствует вариабельность выборки, что обуславливает различные результаты оценки эффекта лечения. Чем меньше исследование, тем больше будет вариабельность выборки, что является еще одним аргументом в пользу концентрации метаанализов на больших исследованиях (15).

Различия в эффективности лечения в пределах метаанализа могут быть связаны со временем и порядком публикации отдельных исследований.

Rothwell и Robertson (22) изучили 26 метаанализов, основанных на 241 исследовании, и обнаружили, что если исследования в метаанализе были расположены в хронологическом порядке, то в ранних исследованиях значительно чаще наблюдались лучшие результаты лечения в исследуемых группах по сравнению с контрольной группой. Эти различия не зависели от объема исследования. Такое наблюдали в 20 из 26 изученных метаанализов.

Это обусловлено, по-видимому, тем, что на начальных этапах изучения эффективности того или

иного метода лечения вероятность того, что оно будет опубликовано, тем выше, чем лучше результаты исследования. Кроме того, оно будет опубликовано быстрее, нежели исследование с негативными результатами. Однако как только лечение начинают считать эффективным и одобряют, становятся интересными и исследования с негативными результатами. Таким образом, метаанализ, проведенный на ранних этапах публикации результатов исследований, переоценивает положительный эффект лечения.

Можно ли избежать гетерогенности в метаанализе?

Сторонники метаанализа обычно стараются убедить читателя, что использованные данные однородны — с целью обоснования объединения разных данных для получения ответа на интересующий вопрос. Различия в клинических исследованиях неизбежны. Исследования могут затрагивать различные популяции пациентов, и даже популяции, определяемые с помощью одних и тех же критериев, изменяются с течением времени. Действительно, пациенты, вовлеченные в сравнительные исследования, могут принадлежать к одной и той же основной популяции, но даже незначительные различия в критериях диагностики, сопутствующих состояниях, тяжести болезни и возрасте, приведут к образованию значительно отличающихся друг от друга групп. Расхождения в дозах, времени включения в исследование и продолжительности терапии могут также приводить к значительным различиям между исследованиями, включенными в метаанализ. Выбор сопутствующего лечения также может отражаться на результатах. Объединение всей информации, содержащейся в ряде исследований, и получение единого показателя могут чрезмерно упростить крайне сложную проблему (23).

Может ли метаанализ прогнозировать результаты единого большого исследования?

Многие исследования, построенные на метаанализе, действительно предоставляют правильную интерпретацию проведенного лечения или процедуры. Приведенный ниже пример заставляет усомниться в правильности этого утверждения. Метаанализ 1266 пациентов из 7 рандомизированных исследований показал, что внутривенное введение магнезии снижает частоту серьезных нарушений ритма и смерти вследствие инфаркта миокарда (отношение шансов 0,44, 95 % доверительный интервал 0,27-0,71) (5). В последующем исследовании LIMIT-2 с участием 2300 пациентов этот результат был подтвержден (относительное снижение смертности составило 24 %) (24). Однако в исследовании ISIS-4 (6) с участием 58 050 пациентов обнаружено незначительное повышение смертности в группе пациентов, получавших магнезию. В недавно проведенном исследовании

MAGIC 30-дневная смертность оказалась одинаковой в обеих группах лечения (15,2%), однако никакого вреда от применения магнезии не было отмечено в восьми заранее выделенных подгруппах (25).

Эта противоречивость результатов была предварительно объяснена систематическими ошибками, связанными с неопубликованными исследованиями (26).

Группа канадских исследователей проанализировала результаты 12 больших (более 1000 пациентов) рандомизированных контролируемых исследований и результаты 19 ранее опубликованных метаанализов по аналогичным темам (23). В 35 % случаев проведение соответствующего метаанализа не помогло точно предсказать результаты больших исследований. Если бы не было последующего рандомизированного исследования, результаты метаанализа привели бы к одобрению неэффективного лечения в 32 % случаев или к отклонению успешного метода лечения в 33 % случаев. Аргентинские исследователи оценили способность метаанализа прогнозировать результаты единого большого исследования (более 1000 пациентов). Было проведено 30 метаанализов, включающих 185 рандомизированных контролируемых исследований. Крупнейшие из этих исследований затем были изъяты и их сравнили с пересчитанным относительным риском метаанализов (27). В 24 из 30 случаев метаанализ показал хорошую прогностическую способность в отношении эффективности лечения в большом исследовании. Однако в остальных 6 случаях (20 %) метаанализ не смог предсказать результат большого исследования, что вызывает сомнения относительно возможности использования такого подхода в каком-либо другом качестве, кроме как инструмента для выдвижения гипотез. Метаанализ имеет тенденцию демонстрировать более выраженные защитные эффекты, нежели проведенное большое исследование.

Кумулятивный метаанализ

Каждый раз, когда докладывают о соответствующем исследовании, свидетельства об эффективности проведенного воздействия накапливаются. В кумулятивном метаанализе предшествующие метаанализы повторяются каждый раз при публикации результатов нового рандомизированного контролируемого исследования (28, 29). Основной целью этого процесса является определение успешности воздействия как можно раньше. Эта корректировка не должна быть непрерывной, а скорее, периодической, например, когда пополнение новой информацией составляет, по крайней мере, 20 % запланированного оптимального объема информации.

Был проведен кумулятивный метаанализ контролируемых исследований бета-блокаторов при вторичной профилактике после инфаркта миокарда (рис. 1) (30). По итогам 13 исследований, опубли-

ликованных к концу 1981 года, относительный риск смертности у пациентов, леченных бета-блокаторами по сравнению с плацебо, составил 0,78 (95 % доверительный интервал 0,69-0,88, $p < 0,001$). Последующие исследования 15 000 пациентов дополнительно подтвердили этот результат, и их можно считать излишними (31). Однако, несмотря на отсутствие общепринятых рекомендаций по интерпретации и достоверности таких повторных анализов, были предложены принципы определения оптимального объема информации и принципы проведения стандартного наблюдения (20).

Закключение

Если предположить, что проблемы внутренней и внешней непротиворечивости отсутствуют, а совокупная оценка имеет клиническое значение, то можно будет сделать следующие выводы:

1. Если метаанализ нескольких исследований, из которых одно или более достигли статистической значимости, привел к положительным результатам, то можно считать, что метаанализ

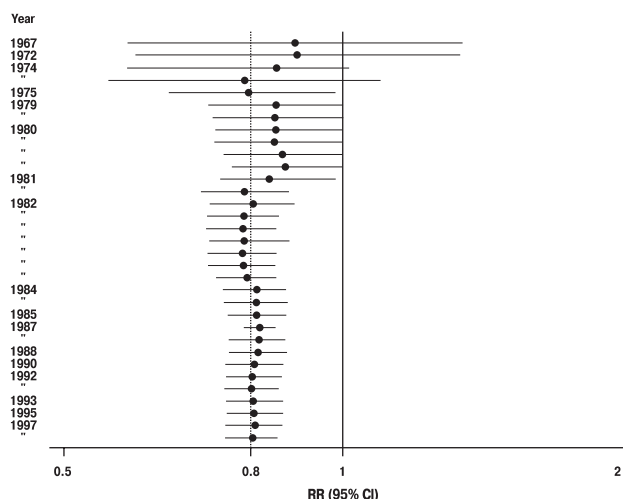


Рис. 1. Кумулятивный метаанализ контролируемых исследований бета-блокаторов при вторичной профилактике после инфаркта миокарда. Отчетливое снижение смертности ($p < 0,001$) стало очевидным с 1981 года (Freemantle et al. Brit. Med. J., 1999, 318, 1730-1737, ref. 30)

укрепил имеющиеся доказательства этого факта.

2. В случаях, когда метаанализ нескольких исследований дал положительный результат, однако при этом не было получено статистической достоверности, следует считать нецелесообразным рекомендовать лечение, основываясь только лишь на данных метаанализа. В подобных случаях, вероятно, требуется провести более крупные исследования.
3. Если метаанализ нескольких исследований, из которых одно или более достигли статистической достоверности, дал статистически достоверный общий результат, следует пересмотреть критерии отбора для исследований и/или

провести дальнейшие клинические исследования.

Таким образом, положительные результаты метаанализа сами по себе не позволяют давать абсолютных рекомендаций по лечению, необходимы и другие доказательства.

Литература

1. Teagarden J.R. Meta-analysis: whether narrative review? Review. Pharmacotherapy, 1989, 9, 274-281.
2. Mulrow C.D. The medical review article: state of the science. Ann. Intern. Med., 1987, 106, 485-488.
3. Yusuf S., Collins R., MacMahon S., Peto R. Effect of intravenous nitrates on mortality in acute myocardial infarction: an overview of the randomized trials. Lancet, 1988, I, 1088-1092.
4. Gruppo Italiano per lo Studio della Streptochinasi nell'Infarto Miocardico (GISSI). GISSI-3: effects of lisinopril and transdermal glyceryl trinitrate singly and together on 6-week mortality and ventricular function after acute myocardial infarction. Lancet, 1994, 343, 1115-1122.
5. Teo K.K., Yusuf S., Collins R., Held P.H., Peto R. Effects of intravenous magnesium in suspected acute myocardial infarction: overview of randomized trials. Br. Med. J., 1991, 303, 1499-503.
6. ISIS-4. Collaboration Group ISIS-4: a randomized factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium in 58,050 patients with suspected acute myocardial infarction. Lancet, 1995, 345, 669-685.
7. Stuck A.E., Siu A.L., Wieland G.D., Adams J., Rubenstein I.Z. Comprehensive geriatric assessment: a meta-analysis of controlled trials. Lancet, 1993, 342, 1032-1036.
8. Reuben D.B., Borok G.M., Wolde-Tsadik G., Ershoff D.H. Randomized trial of comprehensive geriatric assessment in the care of hospitalized patients. N. Engl. J. Med., 1995, 332, 1345-1350.
9. Imperiale T.E., Petrullis A.S. A meta-analysis of low-dose aspirin for the prevention of pregnancy-induced hypertensive disease. JAMA, 1991, 266, 261-265.
10. CLASP Collaborative Group. CLASP: a randomized trial of low-dose aspirin for the prevention and treatment of pre-eclampsia among 9364 pregnant women. Lancet, 1994, 343, 619-629.
11. Boissel J.P., Blanchard J., Panak E., Peyrieux J.C., Saks H. Considerations for the meta-analysis of randomized clinical trials. Summary of a panel discussion. Contr. Clin. Trials, 1989, 10, 254-281.
12. Juni P., Altman D.G., Egger R.M. Assessing the quality of controlled clinical trials. In Egger M., Smith D.G., Altman D.G. (eds). Systematic reviews in health care: analysis in context. London, BMJ Books, 2001, 87-108.
13. Scherer R.W., Dickersin K., Langenberg P. Full publications of results initially presented in abstracts. A meta-analysis. JAMA, 1994, 272, 158-162.
14. Dickersin K. How important is publication bias? A synthesis of available data. AIDS Educ. Prev., 1997, 9, 15-21.
15. Egger M, Smith G.D., Sterne J.A.C. Uses and abuses of meta-analysis. Clin. Med. JRCPL, 2001, 1, 478-484.

16. Ioannidis J.P.A. Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *JAMA*, 1998, 279, 281-286.
17. Horton R., Smith R. Time to register randomized trials. *Lancet*, 1999, 354, 1130-1139.
18. Chalmers I., Altman D. How can medical journals help prevent poor medical research? Some opportunities presented by electronic publishing. *Lancet*, 1999, 353, 490-493.
19. Chalmers T.C., Levin H., Sacks H.S., Reitman D., Berrier J., Nagalingan R. Meta-analysis of clinical trials as a scientific discipline. I: control of bias and comparison with large co-operative trials. *Stat. Med.*, 1987, 6, 315-325.
20. Pogue J., Yusuf S. Overcoming the limitations of current meta-analysis of randomized controlled trials. *Lancet*, 1998, 351, 47-52.
21. Victor N. The challenge of meta-analysis: indications and contra-indications for meta-analysis. *J. Clin. Epidemiol.*, 1995, 48, 133-146.
22. Rothwell P.M., Robertson G. Meta-analysis of randomized controlled trials. *Lancet*, 1997, 350, 1181-1182.
23. Le Lorier J., Gregoire G., Benhaddad A., Lapierre J., Derderian F. Discrepancy between meta-analyses and subsequent large randomized, controlled trials. *N. Engl. J. Med.*, 1997, 337, 536-541.
24. Woods K.L., Fletcher S., Roffe C., Haider Y. Intravenous magnesium sulphate in suspected myocardial infarction: results of the Second Leicester Intravenous Magnesium Intervention Trial (LIMIT-2). *Lancet*, 1992, 239, 15553-15558.
25. The Magnesium in Coronaries Trial Investigators. Early administration of intravenous magnesium to high-risk patients with acute myocardial infarction in the Magnesium in Coronaries (MAGIC) Trial: a randomized control trial. *Lancet*, 2002, 360, 1189- 1196.
26. Parmley W.W. Publication bias. *JACC*, 1994, 24, 1424-430.
27. Villar J., Carroli G., Belizan J.M. Predictive ability of meta-analyses of randomized controlled trials. *Lancet*, 1995, 345, 772-776.
28. Lau J., Antman E.M., Jimenez-Silva J., Kupelnick B., Mosteller F., Chalmers T.C. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N. Engl. J. Med.*, 1992, 327, 248-254.
29. Chalmers I. The Cochrane collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. *Ann. NY Acad. Sci.*, 1993, 703, 156-165.
30. Freemantle N., Cleland J., Young P., Mason J., Harrison J. Beta-blockade after myocardial infarction: systemic review and metaregression analysis. *Brit. Med. J.*, 1999, 318, 1730-1737.
31. Murphy D.J., Povar G.J., Pawlson L.G. Setting limits in clinical medicine. *Arch. Intern. Med.*, 1994, 154, 505-512.