過去の経験に学ぶ，メタアナリシスの本当のエビデンスレベル

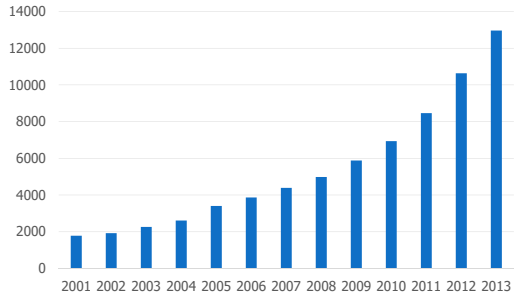
野間 久史
統計数理研究所
2014年12月4日
第35回日本臨床薬理学会学術総会
e-mail: noma@ism.ac.jp
URL: http://www.ism.ac.jp/~noma/

Meta-Analysis

- An analysis of the summary results from two or more similar studies. (Strictly, analyses of analyses; metadata.)
- Such methods are becoming more common and are used as a way of synthesising data from a variety of studies to try to get better answers to specific medical questions.

Dictionary for Clinical Trials, Simon Day, 1999.

Number of Publications



PubMed search: "meta-analysis"

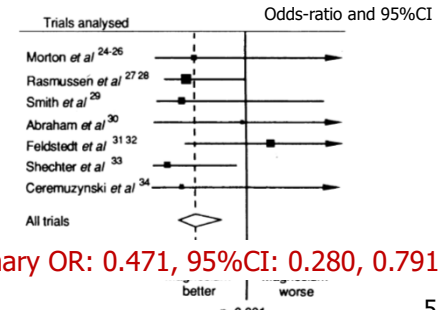
Evidenceレベルの分類

レベル	治療/予防, 病因/害
1a	RCTのシステマティックレビュー（homogeneityであるもの*）
1b	個々のRCT（信頼区間が狭いもの†）
1c	悉無研究（all or none）§
2a	コホート研究のシステマティックレビュー（homogeneity*であるもの）
2b	個々のコホート研究（質の低いRCTを含む；（例）フォローアップ80%未満）
2c	「アウトカム」研究；エコロジー研究
3a	ケースコントロール研究のシステマティックレビュー（homogeneity*であるもの）
3b	個々のケースコントロール研究
4	症例集積研究（および質の低いコホート研究あるいはケースコントロール研究‖§）
5	系統的な批判的吟味を受けていない，または生理学や基礎実験，原理に基づく専門家の意見

http://www.grade-jpn.com/sakusaku/OxCEBM-tx.htm

Oxford Centre for Evidence-Based Medicine Levels of Evidence (2009)

急性心筋梗塞とマグネシウム静脈内投与



Summary OR: 0.471, 95%CI: 0.280, 0.791

Teo et al. (1991)

後続の大規模RCTの結果

- Teo et al. (1991) によるメタアナリシス
  - OR: 0.471, 95%CI: 0.280, 0.791
- ISIS-4 (4th International Study of Infarct Survival, N=58,050)
  - OR: 1.06, 95%CI: 1.00, 1.12
- 後続の大規模ランダム化比較試験によって、まったく結論が覆ってしまった!?

ISIS-4 Collaborative Group (1995)

The New England Journal of Medicine

Special Article

DISCREPANCIES BETWEEN META-ANALYSES AND SUBSEQUENT LARGE RANDOMIZED, CONTROLLED TRIALS

JACQUES LELORIER, M.D., PH.D., GENEVIÈVE GRÉGOIRE, M.D., ABDELTIF BENHADDAD, M.D., JULIE LAPIERRE, M.D., AND FRANÇOIS DERDERIAN, M.SC.

ABSTRACT

Background Meta-analyses are now widely used to provide evidence to support clinical strategies. However, large randomized, controlled trials are considered the gold standard in evaluating the efficacy of clinical interventions.

Methods We compared the results of large randomized, controlled trials (involving 1000 patients or more) that were published in four journals (the *New England Journal of Medicine*, the *Lancet*, the *Annals of Internal Medicine*, and the *Journal of the American Medical Association*) with the results of meta-analyses published earlier on the same topics. Regarding the principal and secondary outcomes, we judged whether the findings of the randomized trials agreed with those of the corresponding meta-analyses, and we determined whether the study results were positive (indicating that treatment improved

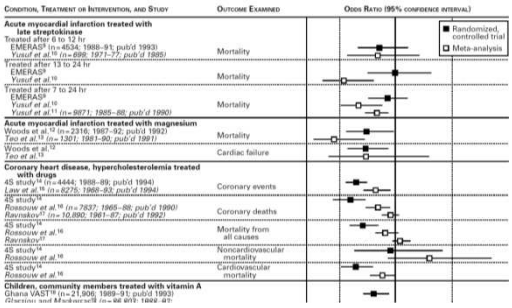
LARGE randomized, controlled trials are generally considered the gold standard in evaluations of the efficacy of clinical interventions. However, since such trials are not always available, clinicians increasingly rely on meta-analysis to support their choice of clinical strategies. Critics have emphasized the intrinsic weaknesses of meta-analysis.[1-5] Pooled results incorporate the biases of individual studies and embody new sources of bias, mostly because of the selection of studies and the inevitable heterogeneity among them.

Although much has been said about the strengths and weaknesses of meta-analysis, there are limited data systematically comparing the results of meta-analyses of several small trials with those of large randomized, controlled trials. Villar et al.[6] reviewed 30 meta-analyses of various interventions in perina-

同様の不一致は、他にも多くの事例で認められていた!!

N Engl J Med 1997;337:536-42.

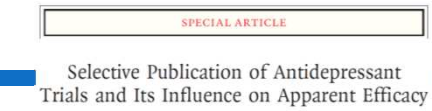
LeLorierらによる報告（一部）



N Engl J Med 1997;337:536-42.

Criticisms of Meta-analysis [1]

- ネガティブな結果が出た試験は報告（出版）されにくく、ポジティブな結果が出た試験は出版されやすい
- 論文として出版された臨床試験の結果を系統的に集めて統合を行うと、ポジティブな方向に偏ったバイアスが入る
- Publication Bias, Reporting Bias



SPECIAL ARTICLE

Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy

Erick H. Turner, M.D., Annette M. Matthews, M.D., Eftihia Linardatos, B.S., Robert A. Tell, L.C.S.W., and Robert Rosenthal, Ph.D.

ABSTRACT

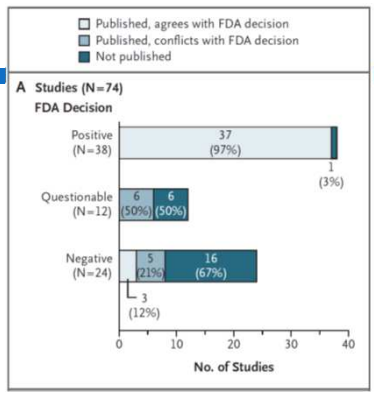
BACKGROUND

Evidence-based medicine is valuable to the extent that the evidence base is complete and unbiased. Selective publication of clinical trials — and the outcomes within those trials — can lead to unrealistic estimates of drug effectiveness and alter the apparent risk–benefit ratio.

METHODS

We obtained reviews from the Food and Drug Administration (FDA) for studies of 12 antidepressant agents involving 12,564 patients. We conducted a systematic literature search to identify matching publications. For trials that were reported in the literature, we compared the published outcomes with the FDA outcomes. We also compared the effect size derived from the published reports with the effect size derived from the entire FDA data set.

N Engl J Med 2008;358:252-60.



N Engl J Med 2008;358:252-60.

Criticisms of Meta-analysis [2]

- 統合される試験の間で、選択基準・除外基準に違いはないか？
- 試験薬の用量，対照薬，アウトカムの定義はすべて同じだろうか？
- 参加した施設の背景はすべての試験で同じだろうか？あるいは、国や地域が異なるのであれば、医療環境に差はないだろうか？
- Mixing Apples and Oranges!!

Criticisms of Meta-analysis [3]

- 系統的なレビューといっても、個々の試験の質が低ければ、統合して得られる情報の質も低い
- ポジティブな結果に偏った小規模な試験のみが対象になっていないか？
- バイアスによって歪められた結果であれば、統合した結果の妥当性も担保されない
- Garbage in, Garbage out!!

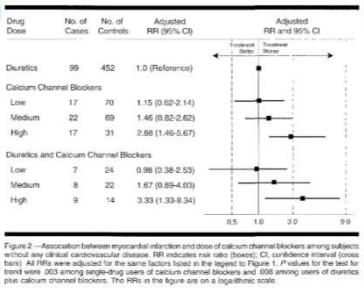
13

Alvan R. Feinstein博士の批評

- メタアナリシスは21世紀における統計学的な錬金術（statistical alchemy）である
- "Significance" can be attained statistically when small group sizes are pooled into big ones.
- We lose or eliminate the elemental scientific requirements for reproducibility and precision, for suitable extrapolation, and even sometimes for fair comparison.

J Clin Epidemiol 1995;48:71-79. 14

Psaty et al. (1995). The risk of myocardial infarction associated with antihypertensive drug therapies. JAMA 274(8):620-625.



結果の再現性が問題とされてきたのは、なにもメタアナリシスだけではない!!

15

間違いはどこに??

- "There is nothing sinful about going out and getting evidence, like asking people how much do you drink and checking breast cancer records. There's nothing sinful about seeing if that evidence correlates. There's nothing sinful about checking for confounding variables. The sin comes in believing a causal hypothesis is true because your study came up with a positive result, or believing the opposite because your study was negative." (Prof. Sander Greenland)

Taubes and Mann (1995) 16

Meta-Analysis

- An analysis of the summary results from two or more similar studies. (Strictly, analyses of analyses; metadata.)
- Such methods are becoming more common and are used as a way of synthesising data from a variety of studies to try to get better answers to specific medical questions.

Dictionary for Clinical Trials, Simon Day, 1999.

17

批評的な評価の「方法」

- 「批評的な評価をするべき」と言っても、広く一般の読者・利用者に使うことができるような、具体的かつ質の高い、標準化されたシステムと方法論が重要
- この20年ほどをかけて（現在進行形で）、これらの方法論は大きく発展している

18

BMJ

BMJ 2011;343:d5928 doi: 10.1136/bmj.d5928 Page 1 of 9

RESEARCH METHODS & REPORTING

The Cochrane Collaboration's tool for assessing risk of bias in randomised trials

OPEN ACCESS

Flaws in the design, conduct, analysis, and reporting of randomised trials can cause the effect of an intervention to be underestimated or overestimated. The Cochrane Collaboration's tool for assessing risk of bias aims to make the process clearer and more accurate

Julian P T Higgins *senior statistician*[1], Douglas G Altman *director*[2], Peter C Gøtzsche *director*[3], Peter Jüni *head of division*[4], David Moher *senior scientist*[5,6], Andrew D Oxman *senior researcher*[7], Jelena Savović *postdoctoral fellow*[8], Kenneth F Schulz *vice president*[9], Laura Weeks *research associate*[3], Jonathan A C Sterne *professor of medical statistics and epidemiology*[8], Cochrane Bias Methods Group, Cochrane Statistical Methods Group

[1]MRC Biostatistics Unit, Institute of Public Health, Cambridge CB2 0SR, UK; [2]Centre for Statistics in Medicine, University of Oxford, Oxford, UK; [3]The Nordic Cochrane Centre, Rigshospitalet and University of Copenhagen, Denmark; [4]Institute of Social and Preventive Medicine, University of Bern, Switzerland; [5]Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada; [6]Department of Epidemiology and Community Medicine, Faculty of Medicine, University of Ottawa, Canada; [7]Preventive and International Health Care Unit, Norwegian Knowledge Centre for the Health Services, Oslo, Norway; [8]Department of Social Medicine, University of Bristol, Bristol, UK; [9]FHI, Research Triangle Park, North Carolina, USA

BMJ 2011;343:d5928.

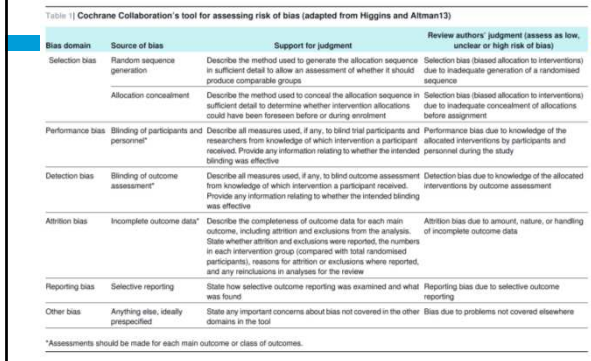
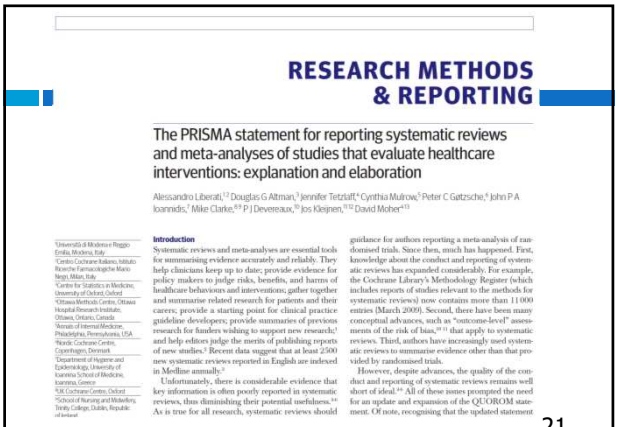


Table 1| **Cochrane Collaboration's tool for assessing risk of bias (adapted from Higgins and Altman13)**

Bias domain	Source of bias	Support for judgment	Review authors' judgment (assess as low, unclear or high risk of bias)
Selection bias	Random sequence generation	Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups	Selection bias (biased allocation to interventions) due to inadequate generation of a randomised sequence
	Allocation concealment	Describe the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen before or during enrolment	Selection bias (biased allocation to interventions) due to inadequate concealment of allocations before assignment
Performance bias	Blinding of participants and personnel*	Describe all measures used, if any, to blind trial participants and researchers from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective	Performance bias due to knowledge of the allocated interventions by participants and personnel during the study
Detection bias	Blinding of outcome assessment*	Describe all measures used, if any, to blind outcome assessment from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective	Detection bias due to knowledge of the allocated interventions by outcome assessment
Attrition bias	Incomplete outcome data*	Describe the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. State whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomised participants), reasons for attrition or exclusions where reported, and any reinclusions in analyses for the review	Attrition bias due to amount, nature, or handling of incomplete outcome data
Reporting bias	Selective reporting	State how selective outcome reporting was examined and what was found	Reporting bias due to selective outcome reporting
Other bias	Anything else, ideally prespecified	State any important concerns about bias not covered in the other domains in the tool	Bias due to problems not covered elsewhere

*Assessments should be made for each main outcome or class of outcomes.

BMJ 2011;343:d5928.



RESEARCH METHODS & REPORTING

The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration

Alessandro Liberati,[1,2] Douglas G Altman,[3] Jennifer Tetzlaff,[4] Cynthia Mulrow,[5] Peter C Gøtzsche,[6] John P A Ioannidis,[7] Mike Clarke,[8,9] P J Devereaux,[10] Jos Kleijnen,[11,12] David Moher[4,13]

Introduction

Systematic reviews and meta-analyses are essential tools for summarising evidence accurately and reliably. They help clinicians keep up to date; provide evidence for policy makers to judge risks, benefits, and harms of healthcare behaviours and interventions; gather together and summarise related research for patients and their carers; provide a starting point for clinical practice guideline developers; provide summaries of previous research for funders wishing to support new research; and help editors judge the merits of publishing reports of new studies.[1] Recent data suggest that at least 2500 new systematic reviews reported in English are indexed in Medline annually.[2]

Unfortunately, there is considerable evidence that key information is often poorly reported in systematic reviews, thus diminishing their potential usefulness.[3-6] As is true for all research, systematic reviews should guidance for authors reporting a meta-analysis of randomised trials. Since then, much has happened. First, knowledge about the conduct and reporting of systematic reviews has expanded considerably. For example, the Cochrane Library's Methodology Register (which includes reports of studies relevant to the methods for systematic reviews) now contains more than 11 000 entries (March 2009). Second, there have been many conceptual advances, such as "outcome-level" assessments of the risk of bias,[9-11] that apply to systematic reviews. Third, authors have increasingly used systematic reviews to summarise evidence other than that provided by randomised trials.

However, despite advances, the quality of the conduct and reporting of systematic reviews remains well short of ideal.[4,6] All of these issues prompted the need for an update and expansion of the QUOROM statement. Of note, recognising that the updated statement

BMJ 2009;339:b2700.

まとめ

- メタアナリシスは統計学的な錬金術ではない
- それを防ぐための科学的な方法論の基盤も、進歩し続けている
- 「有意な結果が出た」という報告を受けて、盲目的に「効く」と信じることこそが誤り
- 正しく批評的な解釈をするために、出版された情報の消費者の側もリテラシーを持つことが重要

参考文献

- Feinstein, A. R. (1995). Meta-analysis: statistical alchemy for the 21st century. Journal of Clinical Epidemiology 48, 71-79.
- ISIS-4 Collaborative Group. (1995). ISIS-4: A randomised factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium sulphate in 58050 patients with suspected acute myocardial infarction. Lancet 345, 669-685.
- Higgins, J. P., Altman, D. G., Gotzsche, P. C., et al. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 343, d5928.
- LeLorier, J., Gregoire, G., Benhaddad, A., Lapierre, J., and Derderian, F. (1997). Discrepancies between meta-analyses and subsequent large randomized, controlled trials. New England Journal of Medicine 337, 536-542.

- Liberati, A., Altman, D. G., Tetzlaff, J., et al. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 339, b2700.
- Oxford Centre for Evidence-Based Medicine Levels of Evidence. (2009). Levels of evidence and grades of recommendations. [http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/]
- Psaty, B. M., Heckbert, S. R., Koepsell, T. D., et al. (1995). The risk of myocardial infarction associated with antihypertensive drug therapies. JAMA 274, 620-625.
- Simon, D. (1999). Dictionary for Clinical Trials. Chichester: John Wiley & Sons.

- Taubes, G., and Mann, C. C. (1995). Epidemiology faces its limits. Science 269, 164-169.
- Teo, K. K., Yusuf, S., Collins, R., Held, P. H., and Peto, R. (1991). Effects of intravenous magnesium in suspected acute myocardial infarction: overview of randomized trials. British Medical Journal 303, 1499-1503.
- Turner, E. H., Matthews, A. M., Linardatos, E., Tell, R., and Rosenthal, R. (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. New England Journal of Medicine 358, 252-260.

25