ネットワークメタアナリシス
ーComparative Effectiveness Researchにおける エビデンス統合の方法ー

野間 久史
情報・システム研究機構　統計数理研究所
2018年1月29日
国立がん研究センター 生物統計セミナー【発展編】
e-mail: noma@ism.ac.jp
URL: http://normanh.skr.jp

営利目的でのご利用や無断転載はご遠慮ください

http://www.ism.ac.jp/

営利目的でのご利用や無断転載はご遠慮ください

July 6, 2015 7:00 am JST

Generic drugs

Japan becomes a new frontier

Teva Pharmaceutical Industries is boosting production capacity for generics to be sold in Japan.

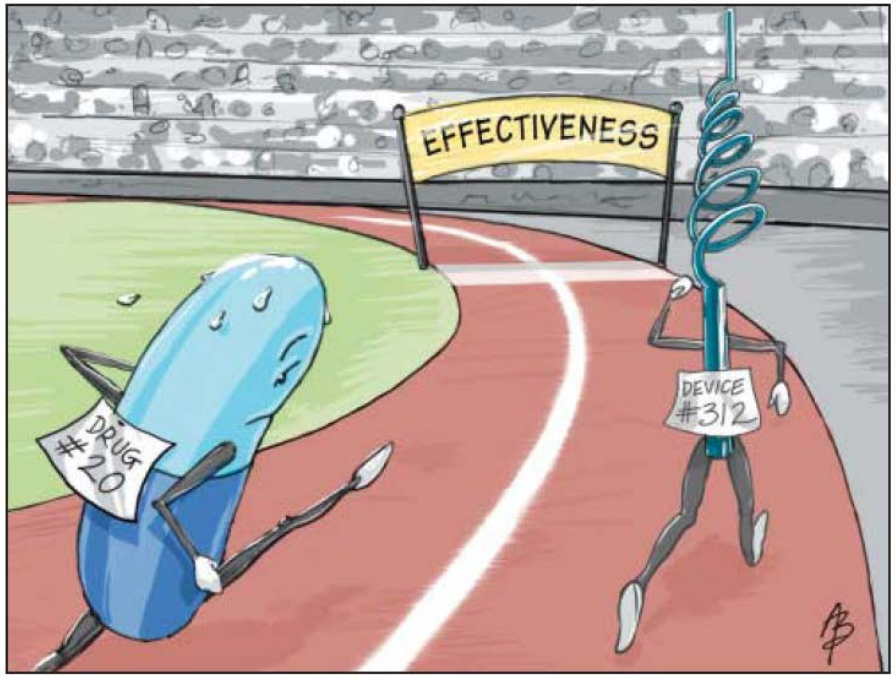
TOKYO -- Leading generic-drug makers are ready to tap into the Japanese market. The government in Tokyo recently moved to promote more use of generics in the country. It aims to raise the share of generic drug prescriptions to 80% between fiscal 2018 and 2020. The policy may change the power dynamics among foreign and Japanese drugmakers.

- 65歳以上人口は25%を超えており、WHOが定める「超高齢社会」に
- 国民医療費は、年間40兆円を超えている（2013年度以降）
- 厚生労働省は、ジェネリック医薬品の使用促進を進めており、平成30年度から平成32年度末までの間のなるべく早い時期にその使用率を80%以上とする数値目標を掲げている

http://asia.nikkei.com/Business/Trends/Japan-becomes-a-new-frontier

営利目的でのご利用や無断転載はご遠慮ください

Comparative Effectiveness Research (CER)



Comparative effectiveness research allows investigators to determine the superiority, inferiority, or equivalence of various interventions when pitted against each other.

Mitka, M. (2010). US Government Kicks Off Program for Comparative Effectiveness Researches. JAMA 304: 2230-1.

- The Institute of Medicine committee has defined CER as “the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition or to improve the delivery of care. The purpose of CER is to assist consumers, clinicians, purchasers, and policy makers to make informed decisions that will improve health care at both the individual and population levels.”

http://www.nap.edu/read/12648/chapter/4

営利目的でのご利用や無断転載はご遠慮ください

ORIGINAL CONTRIBUTION JAMA-EXPRESS

Major Outcomes in High-Risk Hypertensive Patients Randomized to Angiotensin-Converting Enzyme Inhibitor or Calcium Channel Blocker vs Diuretic

The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT)

The ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group

TREATMENT AND COMPLICATIONS among the 50 to 60 million people in the United States with hypertension are estimated to cost $37 billion annually, with antihypertensive drug costs alone accounting for an estimated $15.5 billion per year.[1] Antihypertensive drug therapy substantially reduces the risk of hypertension-related morbidity and mortality.[2-6] However, the optimal choice for initial pharmacotherapy of hypertension is uncertain.[7]

Earlier clinical trials documented the benefit of lowering blood pressure (BP) using primarily thiazide diuretics or β-blockers.[2,3,8] After these studies, several newer classes of antihypertensive

Context Antihypertensive therapy is well established to reduce hypertension-related morbidity and mortality, but the optimal first-step therapy is unknown.

Objective To determine whether treatment with a calcium channel blocker or an angiotensin-converting enzyme inhibitor lowers the incidence of coronary heart disease (CHD) or other cardiovascular disease (CVD) events vs treatment with a diuretic.

Design The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT), a randomized, double-blind, active-controlled clinical trial conducted from February 1994 through March 2002.

Setting and Participants A total of 33357 participants aged 55 years or older with hypertension and at least 1 other CHD risk factor from 623 North American centers.

Interventions Participants were randomly assigned to receive chlorthalidone, 12.5 to 25 mg/d (n=15255); amlodipine, 2.5 to 10 mg/d (n=9048); or lisinopril, 10 to 40 mg/d (n=9054) for planned follow-up of approximately 4 to 8 years.

Main Outcome Measures The primary outcome was combined fatal CHD or nonfatal myocardial infarction, analyzed by intent-to-treat. Secondary outcomes were all-cause mortality, stroke, combined CHD (primary outcome, coronary revascularization, or angina with hospitalization), and combined CVD (combined CHD, stroke, treated angina without hospitalization, heart failure [HF], and peripheral arterial disease).

Results Mean follow-up was 4.9 years. The primary outcome occurred in 2956 participants, with no difference between treatments. Compared with chlorthalidone (6-year rate, 11.5%), the relative risks (RRs) were 0.98 (95% CI, 0.90-1.07) for amlo-

JAMA 2002, 288: 2981-2196.

- ALLHAT試験：1990年代、米国で行われた、降圧薬のComparative Effectivenessを比較した臨床試験
- 利尿剤、β-blocker、ACE阻害薬、Ca拮抗薬の4薬剤を比較した試験
- 当初、計画段階で見積もられたサンプルサイズは、4万人！（検定の多重性も考慮しなくてはならない）
- Head-to-headの比較試験で、Comparative Effectivenessの検証的なエビデンスを得るためには、膨大なサンプルサイズ、費用、時間がかかる！

営利目的でのご利用や無断転載はご遠慮ください

Evidence Synthesis in Medical Research

- 米国NCBIが運営する医学・生物学領域の文献データベースPubMedの収録文献数は…
- 2016年の1年間で1,258,755件!!
- 科学的根拠に基づく最善の医療を行うために、日々蓄積・更新される医学研究のエビデンスを、適切に統合し、良質な医療情報を構築するためには、有効なデータ解析の方法論が必要

http://english.stackexchange.com

営利目的でのご利用や無断転載はご遠慮ください

コクラン共同計画　Cochrane Collaboration

http://www.cochrane.org/

- 医学研究において蓄積された最新・最良のエビデンスを批判的にまとめたシステマティック・レビューを作成し、それらの情報を必要とするユーザーへ広く提供することを目的とした国際共同プロジェクト
- 蓄積されたエビデンスを統合し、科学的に妥当な評価を行うためには、適切な統計手法が不可欠
- メタアナリシス（Meta-Analysis）

営利目的でのご利用や無断転載はご遠慮ください

抗うつ薬のメタアナリシス

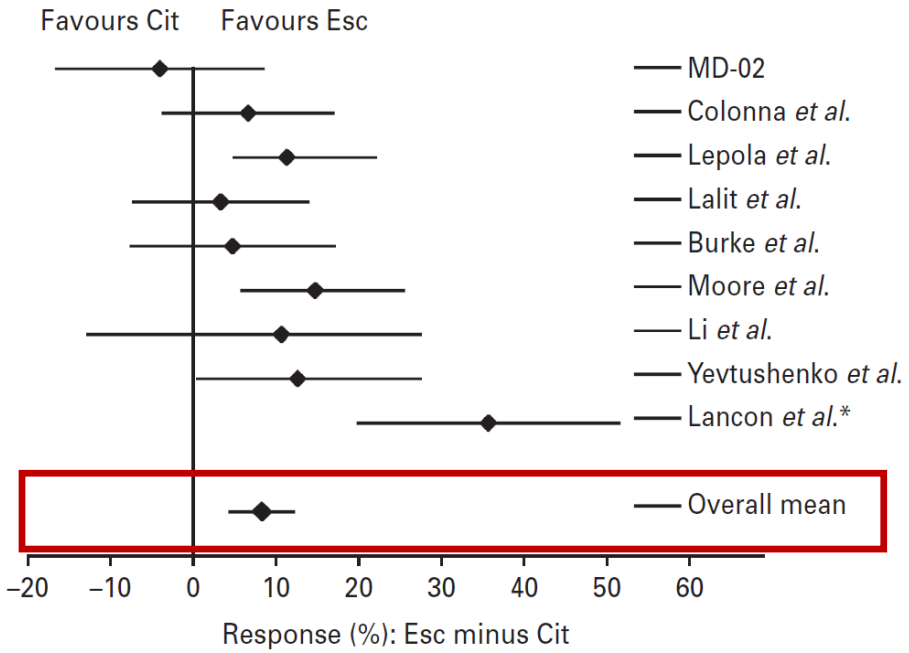


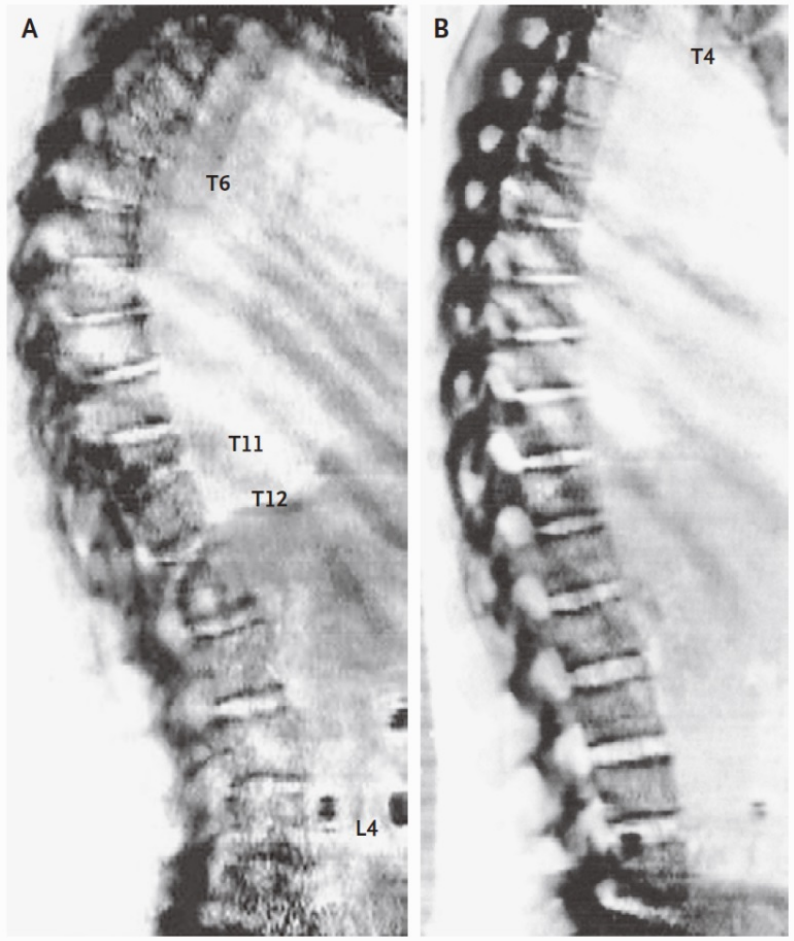
Fig. 2. Estimated difference in response rates at week 8 (or last assessment if <8 wk) (Esc−Cit) with 95% confidence intervals. Esc, Escitalopram; Cit, citalopram. Response is defined as the percentage of patients with ⩾50% improvement from baseline (MADRS or $HAMD_{17}$). * The naturalistic study of Lançon *et al.* (2006) was not included in the calculation of the overall mean.

- EscitaropramとCitaropramの有効性を比較したメタアナリシス
- ESC vs. CITの比較を行った8つのランダム化臨床試験を系統的に収集し、その結果を統合
  - N=2009（ESC 995人, CIT 1014人）
- 8つの試験の結果によって推定された総合的な治療効果は、反応率の差8.3% (95%CI: 4.4-12.3) で、ESCのほうが有意に反応率が高いという結果に

Montgomery et al. (2011)

営利目的でのご利用や無断転載はご遠慮ください

骨粗鬆症　Osteoporosis



- 米国では、1000万人が有病，3400万人にその潜在的なリスクがある
- 1年あたり150万件の脆弱性骨折が発生，医療・介護費用は年間200億ドル
- 本邦においても、大腿骨頸部骨折と脊椎骨折を合わせた総医療費は2382～3218億円，医療・介護費用は7974～9895億円と推定
- 少々古いデータですが、ご容赦ください

Holroyd et al. (2008), Favus (2010), 原田ら (2005)

営利目的でのご利用や無断転載はご遠慮ください

治療の選択肢

- 10種類以上の治療法とその組み合わせ
  - ホルモン療法（エストロゲン補充療法）
  - 選択的エストロゲン受容体モジュレーター (SERM)
  - カルシウム and/or ビタミンD
  - ビスフォスフォネート製剤（alendronate, risedronate など）
  - その他のホルモン製剤 （teriparatideなど）
- Cost：ranges from $4 to $130 per month

Li and Becker (2013)

営利目的でのご利用や無断転載はご遠慮ください

14のCochrane Reviews

- “At 10 mg per day (alendronate), both clinically important and statistically significant reductions in vertebral, non-vertebral, hip and wrist fractures were observed for secondary.” (Wells et al., 2008a)
- “No statistically significant reductions in non-vertebral, hip, or wrist fractures were found, regardless of whether etidronate was used for primary or secondary prevention.” (Wells et al., 2008b)
- “Vitamin D alone appears unlikely to be effective in preventing hip fracture, vertebral fracture or any new fracture. Vitamin D with calcium reduces hip fractures.” (Avenell et al., 2009)
- **Which interventions work? In whom?**

Li and Becker (2013)

営利目的でのご利用や無断転載はご遠慮ください

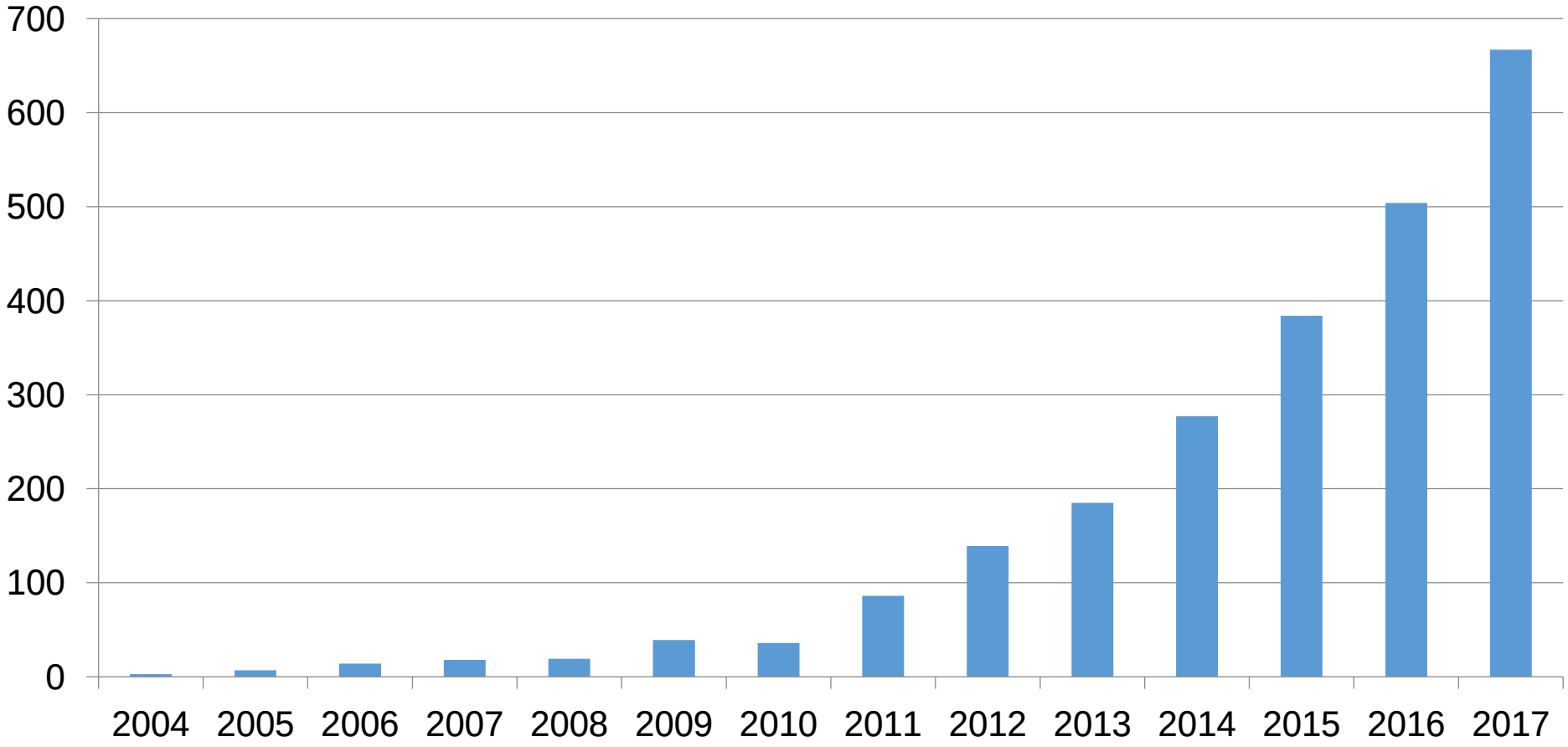
Network Meta-Analysis (NMA)

- 有効性についてのエビデンスが確立された複数の治療法のComparative Effectivenessを評価するために新たに開発されたエビデンス統合のための方法論
- 2種類の治療の対比較に限らず、利用可能な複数（3種類以上）の治療法を、過去の臨床試験から得られたエビデンスを統合し、系統的に比較するためのメタアナリシスの方法論
- Multiple treatment comparison meta-analysis, Mixed-treatment comparisonなどともいわれる

Caldwell et al. (2005), Salanti (2012)

営利目的でのご利用や無断転載はご遠慮ください

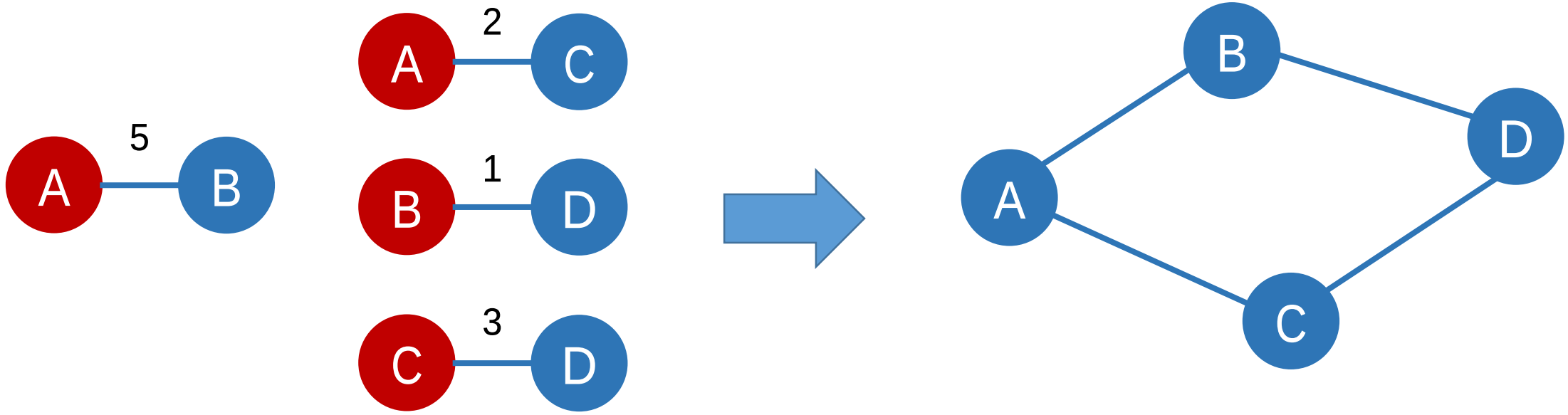
Number of Publications



Pubmed search: ("network meta-analysis" OR "multiple treatment comparison meta-analysis“ OR "mixed comparison meta-analysis")

営利目的でのご利用や無断転載はご遠慮ください

‘Network’ Meta-Analysis



赤：当該試験のActive Treatment
数字：試験数

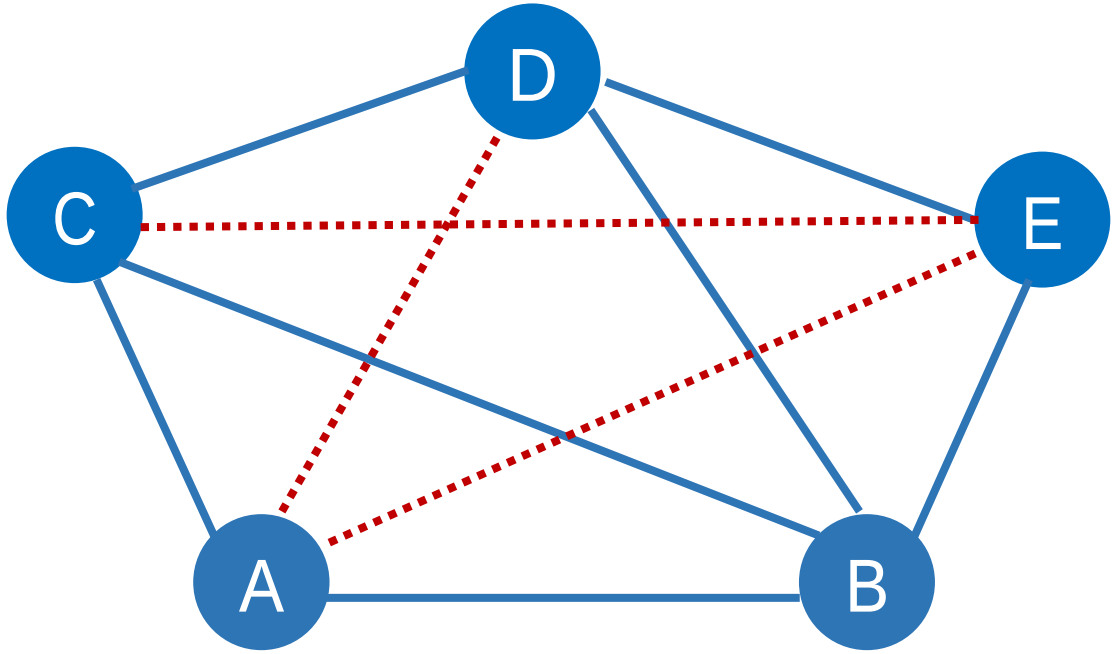
※ 個々の臨床試験では、ほとんどが2つの治療の対比較が行われており、多くても3-4ほどの治療の比較のみ

直接比較のエビデンスのパスをつなげて、治療のネットワークを作り、対象となる治療（A-D）間の治療効果の比較を行う

営利目的でのご利用や無断転載はご遠慮ください

‘Network’ Meta-Analysis

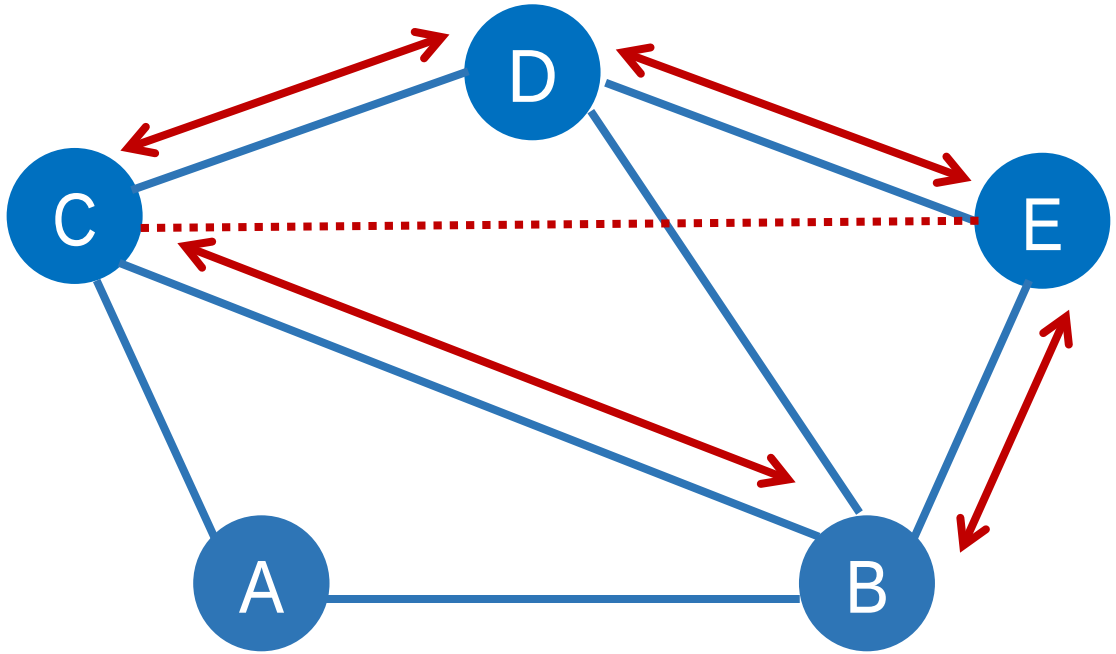
- 直接比較のエビデンスがあるパスを通して、治療効果を比較することができる
- 直接比較のエビデンスがないパスでの比較を行うことはできる？？



営利目的でのご利用や無断転載はご遠慮ください

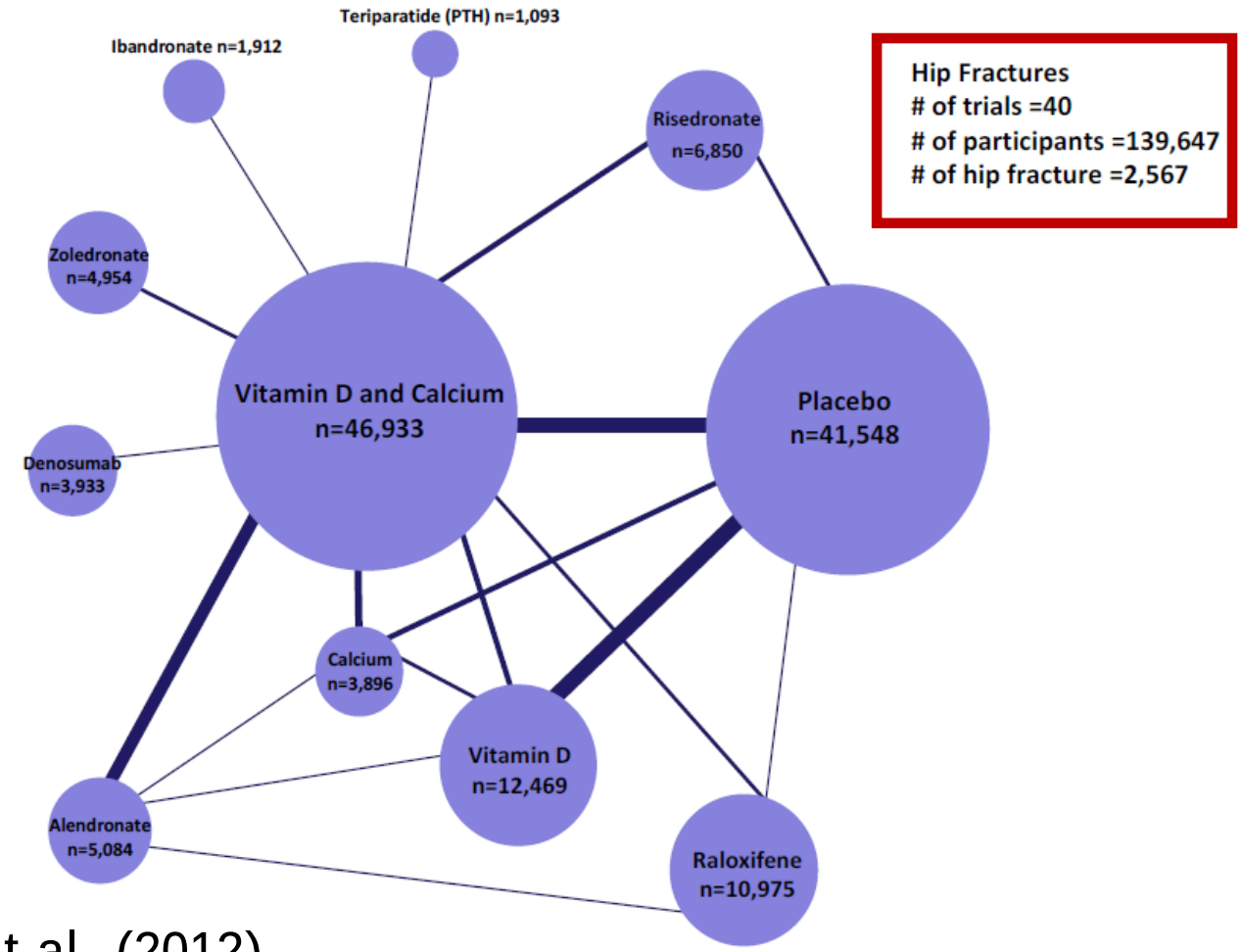
Synthesizing *Indirect* Evidence

- 直接比較のエビデンスだけでは、すべての治療のComparative Effectivenessを評価することができない？？
- ネットワーク上の間接的な連結のあるパスの情報を活用する！！



営利目的でのご利用や無断転載はご遠慮ください

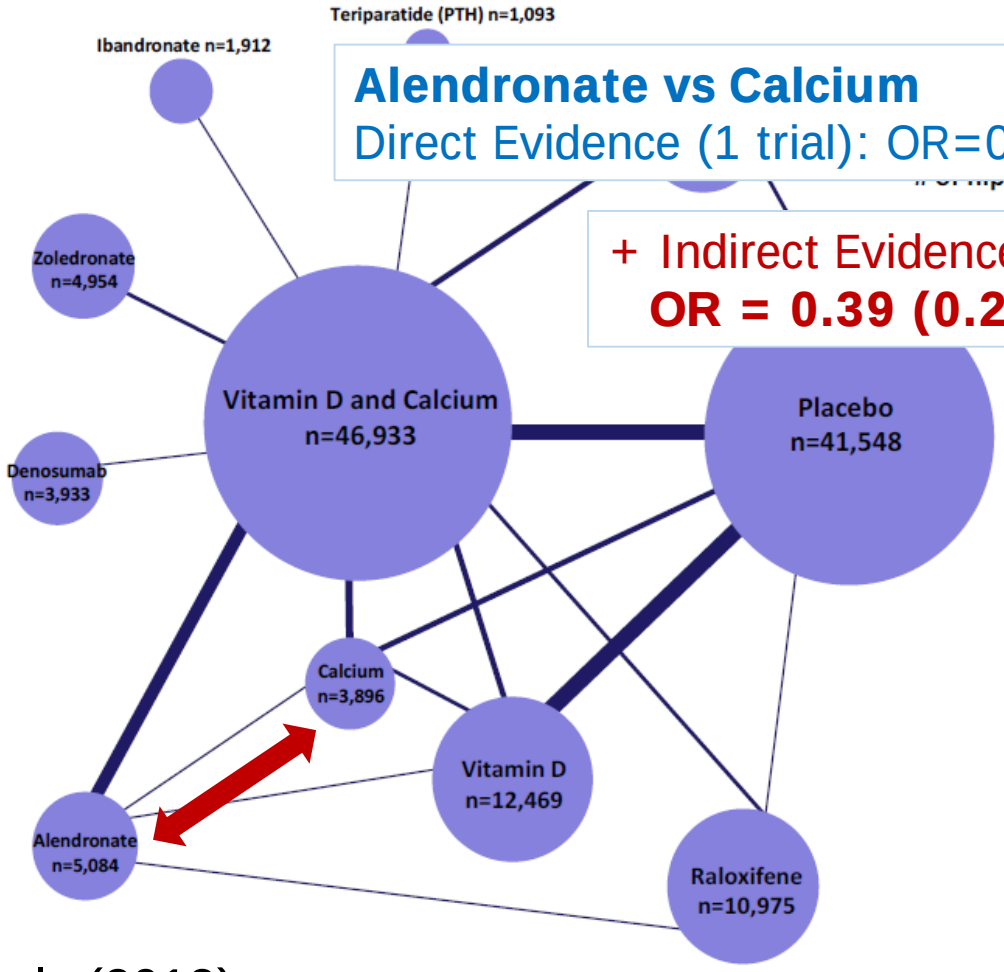
Evidence Network (Hip Fractures)



Murad et al. (2012)

営利目的でのご利用や無断転載はご遠慮ください

Evidence Network (Hip Fractures)



Alendronate vs Calcium
Direct Evidence (1 trial): OR=0.22 (95%CI: 0.02, 2.13)

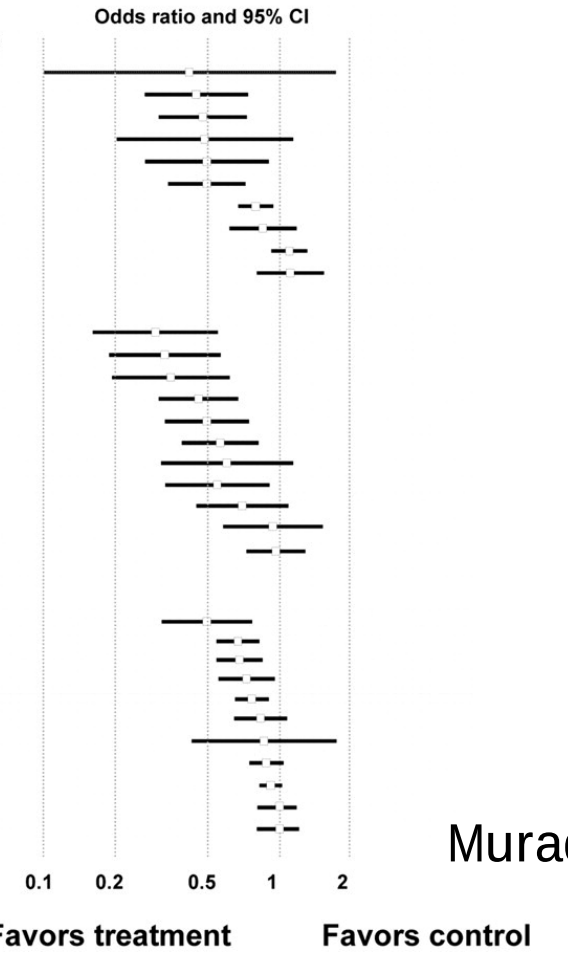
+ Indirect Evidence (Many trials):
OR = 0.39 (0.23, 0.64)

Murad et al. (2012)

営利目的でのご利用や無断転載はご遠慮ください

Comparative Efficacyの評価

		Odds ratio	Lower limit	Upper limit	p-Value
Hip Fractures	Teriparatide	0.42	0.10	1.82	0.24
	Alendronate	0.45	0.27	0.68	0.00
	Risedronate	0.48	0.31	0.66	0.00
	Ibandronate	0.49	0.20	1.82	0.11
	Denosumab	0.50	0.27	0.86	0.03
	Zoledronate	0.50	0.34	0.73	0.00
	Vitamin D+ Calcium	0.81	0.68	0.96	0.02
	Raloxifene	0.87	0.63	1.22	0.41
	Vitamin D	1.13	0.95	1.34	0.18
	Calcium	1.14	0.82	1.59	0.44
Vertebral Fractures	Teriparatide	0.30	0.16	0.55	0.00
	Denosumab	0.33	0.19	0.65	0.00
	Zoledronate	0.35	0.20	0.64	0.00
	Risedronate	0.46	0.31	0.68	0.00
	Alendronate	0.50	0.33	0.79	0.00
	Raloxifene	0.57	0.39	0.83	0.00
	Bazedoxifene	0.61	0.32	1.18	0.14
	Ibandronate	0.62	0.37	0.98	0.04
	Calcium	0.71	0.45	1.12	0.14
	Vitamin D	0.96	0.59	1.58	0.87
	Vitamin D+ Calcium	0.99	0.74	1.41	0.95
Non verterbral Fractures	Teriparatide	0.50	0.32	0.78	0.00
	Risedronate	0.68	0.55	0.81	0.00
	Zoledronate	0.69	0.55	0.84	0.00
	Denosumab	0.74	0.56	0.94	0.03
	Alendronate	0.78	0.66	0.92	0.00
	Bazedoxifene	0.85	0.65	1.11	0.23
	Ibandronate	0.88	0.43	1.64	0.73
	Raloxifene	0.90	0.76	1.03	0.22
	Vitamin D+ Calcium	0.94	0.84	1.02	0.28
	Calcium	1.00	0.82	1.22	1.00
	Vitamin D	1.01	0.82	1.20	0.93



Murad et al. (2012)

※ Baseline treatment: Placebo

営利目的でのご利用や無断転載はご遠慮ください

Home Journals Specialties The Lancet Clinic Global Health Multimedia Campaigns More Information for Submit a Paper

THE LANCET

Login | Register | Subscribe

Online First Current Issue All Issues Special Issues Multimedia Information for Authors

All Content Search Advanced Search

< Previous Article Volume 373, No. 9665, p746–758, 28 February 2009 Next Article >

Access this article on ScienceDirect

Articles

Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis

Dr Andrea Cipriani, PhD, Toshiaki A Furukawa, MD, Georgia Salanti, PhD, John R Geddes, MD, Julian PT Higgins, PhD, Rachel Churchill, PhD, Norio Watanabe, PhD, Atsuo Nakagawa, MD, Ichiro M Omori, PhD, Hugh McGuire, MA, Michele Tansella, MD, Corrado Barbui, MD

Published: 29 January 2009

PlumX Metrics

DOI: http://dx.doi.org/10.1016/S0140-6736(09)60046-5

Article Info

Summary Full Text Tables and Figures References

Summary

Background

Conventional meta-analyses have shown inconsistent results for efficacy of second-generation antidepressants. We therefore did a multiple-treatments meta-analysis, which accounts for both direct and indirect comparisons, to assess the effects of 12 new-generation antidepressants on major depression.

Methods

We systematically reviewed 117 randomised controlled trials (25 928 participants) from 1991 up to Nov 30, 2007, which compared any of the following antidepressants at therapeutic dose range for the acute treatment of unipolar major depression in adults: bupropion, citalopram, duloxetine, escitalopram, fluoxetine, fluvoxamine, milnacipran, mirtazapine, paroxetine, reboxetine, sertraline, and venlafaxine. The main outcomes were the proportion of patients who responded to or dropped out of the allocated treatment. Analysis was done on an intention-to-treat basis.

Findings

Article Options

PDF (253 KB)
Download Images(.ppt)
Email Article
Add to My Reading List
Export Citation
Create Citation Alert
Cited by in Scopus (872)
Request Permissions
Altmetric 85

Linked Articles

COMMENT
Antidepressants are not all created equal

This article can be found in the following collections: Anxiety & depression

Popular Articles

Most Read Most Cited

Most read in *The Lancet* within the past 30 days.

- 12種類の新世代抗うつ薬の有効性・許容性を比較したネットワークメタアナリシス
- 1991年から2007年に実施された、12種類の新世代抗うつ薬のいずれかを比較したランダム化臨床試験（117試験，25,298 人）の結果を統合解析
- 精神医学領域で初めて行われた、最大規模のネットワークメタアナリシス
- 有効性に関するアウトカムは、割りつけられた治療への反応（=0,1）
- 効果の指標には、Odds Ratio (OR) を使用

Cipriani, Furukawa, Salanti *et al*. (2009). *Lancet* **373**: 746-58.

営利目的でのご利用や無断転載はご遠慮ください

12種類の新世代抗うつ薬のNMA

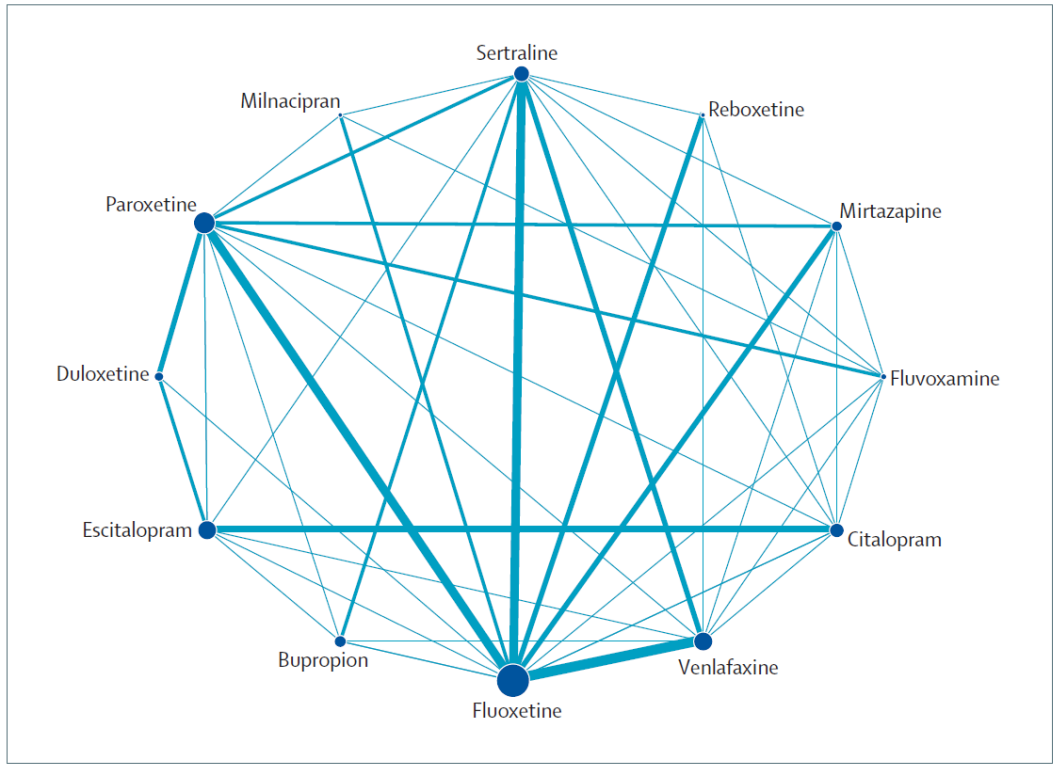


Figure 2: **Network of eligible comparisons for the multiple-treatment meta-analysis for efficacy (response rate)**
The width of the lines is proportional to the number of trials comparing each pair of treatments, and the size of each node is proportional to the number of randomised participants (sample size). The network of eligible comparisons for acceptability (dropout rate) analysis is similar.

- 1991年から2007年に実施された、左記の12種類の抗うつ薬のいずれかを比較したランダム化臨床試験（117試験；25,298 人）を系統的に収集
- 有効性に関するアウトカムは、割りつけられた治療への反応の有無（=0,1）
- ネットワークメタアナリシスによって、12種類の抗うつ薬の有効性・許容性を系統的に評価

Cipriani, Furukawa, Salanti et al. (2009)

営利目的でのご利用や無断転載はご遠慮ください

Study Selection and Data Collection

- For our analysis, we included only randomised controlled trials that compared any of the following 12 new-generation antidepressants (bupropion, citalopram, duloxetine, escitalopram, fluoxetine, fluvoxamine, milnacipran, mirtazapine, paroxetine, reboxetine, sertraline, and venlafaxine) as monotherapy in the acute-phase treatment of adults with unipolar major depression. We excluded placebo groups where present and randomised controlled trials of women with post-partum depression.

Cipriani, Furukawa, Salanti et al. (2009)

営利目的でのご利用や無断転載はご遠慮ください

ICR臨床研究入門 http://www.icrweb.jp

Study Selection and Data Collection

- Two persons within the reviewing team independently reviewed references and abstracts retrieved by the search, assessed the completeness of data abstraction, and confirmed quality rating. We used a structured data abstraction form to ensure consistency of appraisal for each study. Investigators were contacted and asked to provide data to supplement the incomplete reporting of the original articles.
- We gave studies a quality rating of adequate, unclear, or inadequate, according to the adequacy of the random allocation concealment and blinding. Studies that scored adequate or unclear on these criteria were included in the final list.

Cipriani, Furukawa, Salanti et al. (2009)

営利目的でのご利用や無断転載はご遠慮ください

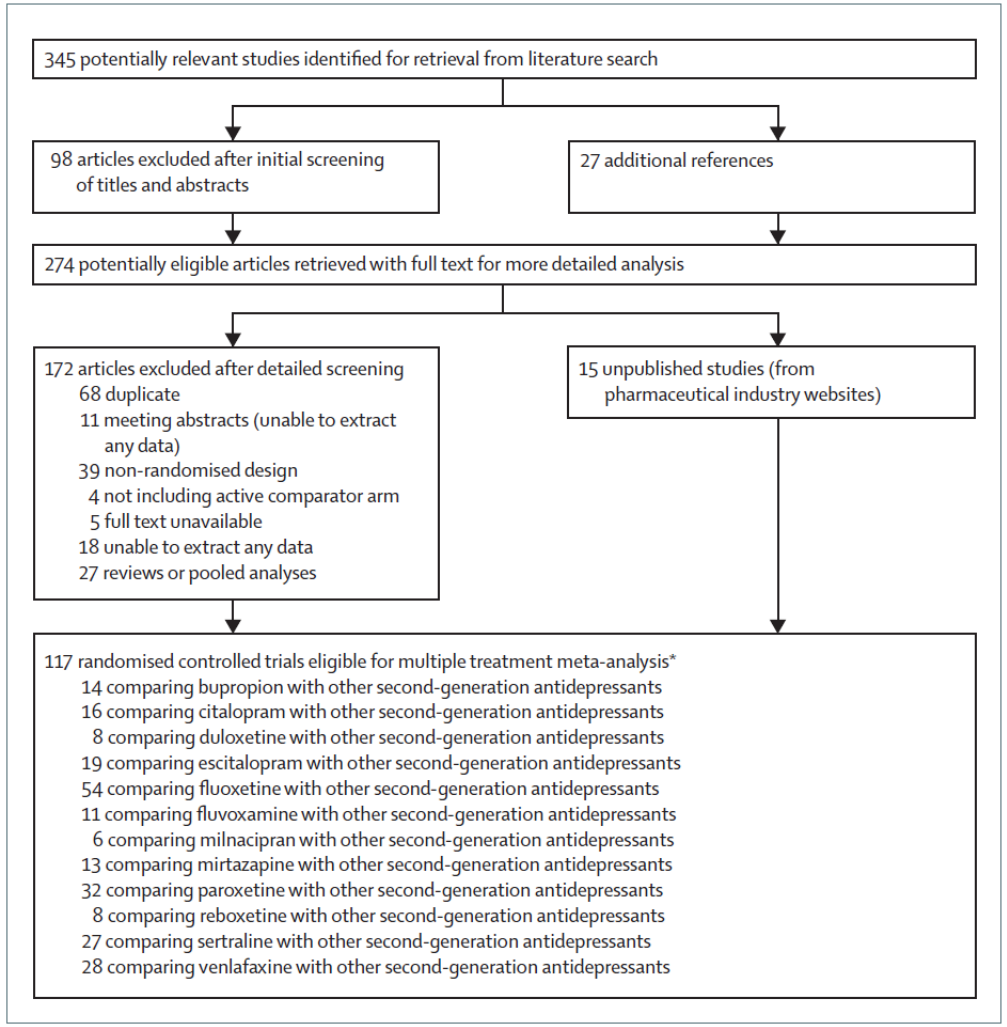


Figure 1: **Study selection process**
*117 randomised controlled trials correspond to 236 arms because two three-arm studies comparing fluoxetine with paroxetine and sertraline were included in this multiple-treatments meta-analysis.

Cipriani, Furukawa, Salanti et al. (2009)

営利目的でのご利用や無断転載はご遠慮ください

	Number of trials	Year of publication			Country				
		Earliest	Median	Latest	Europe	North America	Africa	Asia	Multiple countries
Bupropion	14	1991	2003	2007	1	10	0	0	2
Citalopram	16	1993	2002	2007	4	4	0	1	4
Duloxetine	8	2002	2006	2007	2	5	0	0	1
Escitalopram	19	2000	2005	2007	5	11	0	0	2
Fluoxetine	54	1991	2000	2007	15	13	1	3	6
Fluvoxamine	11	1993	1998	2006	3	2	0	1	2
Milnacipran	6	1994	2000	2003	2	1	0	2	0
Mirtazapine	13	1997	2002	2005	3	3	1	1	5
Paroxetine	32	1993	2001	2007	12	13	1	1	2
Reboxetine	8	1997	2003	2006	2	2	0	0	1
Sertraline	27	1993	2000	2007	10	9	0	2	1
Venlafaxine	28	1994	2002	2007	7	5	0	1	6

The number of studies across countries in this table does not match the number of trials included in the review. Missing studies scored as other or not known. *Two three-arm studies comparing fluoxetine with paroxetine and sertraline were included in the systematic review (the total number of arms is 236 and it corresponds to 115 two-arm and two three-arm studies).

Table 2: **Studies included in the multiple-treatments meta-analysis**

Cipriani, Furukawa, Salanti et al. (2009)

営利目的でのご利用や無断転載はご遠慮ください

Network Plot

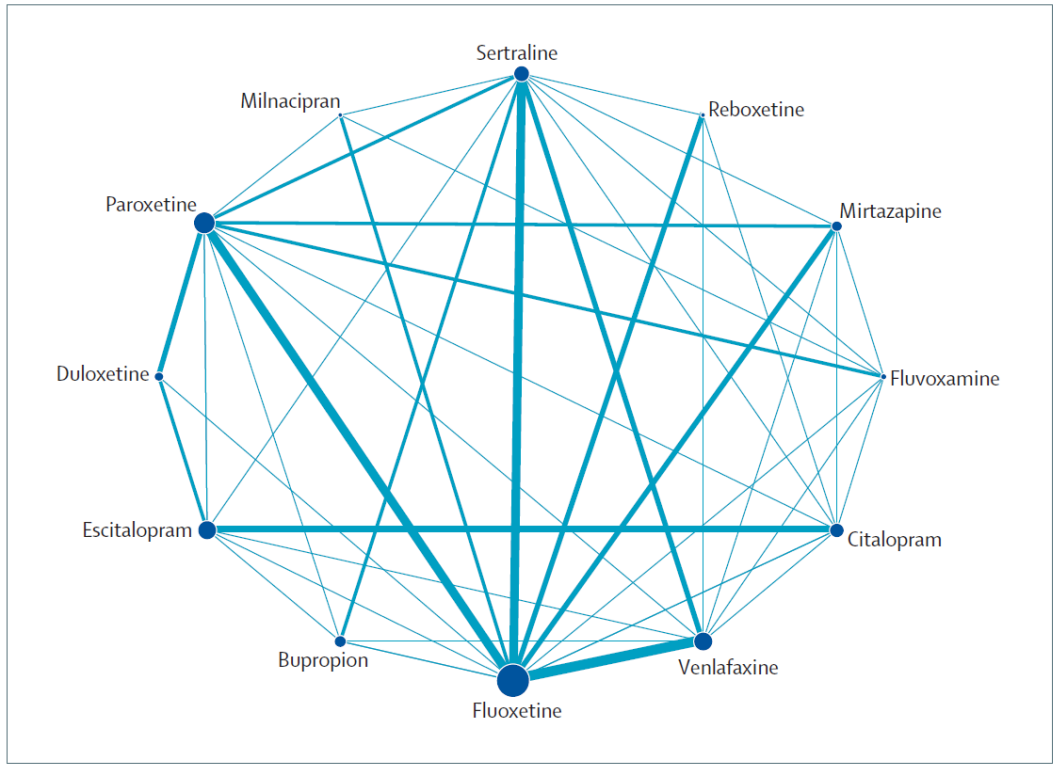


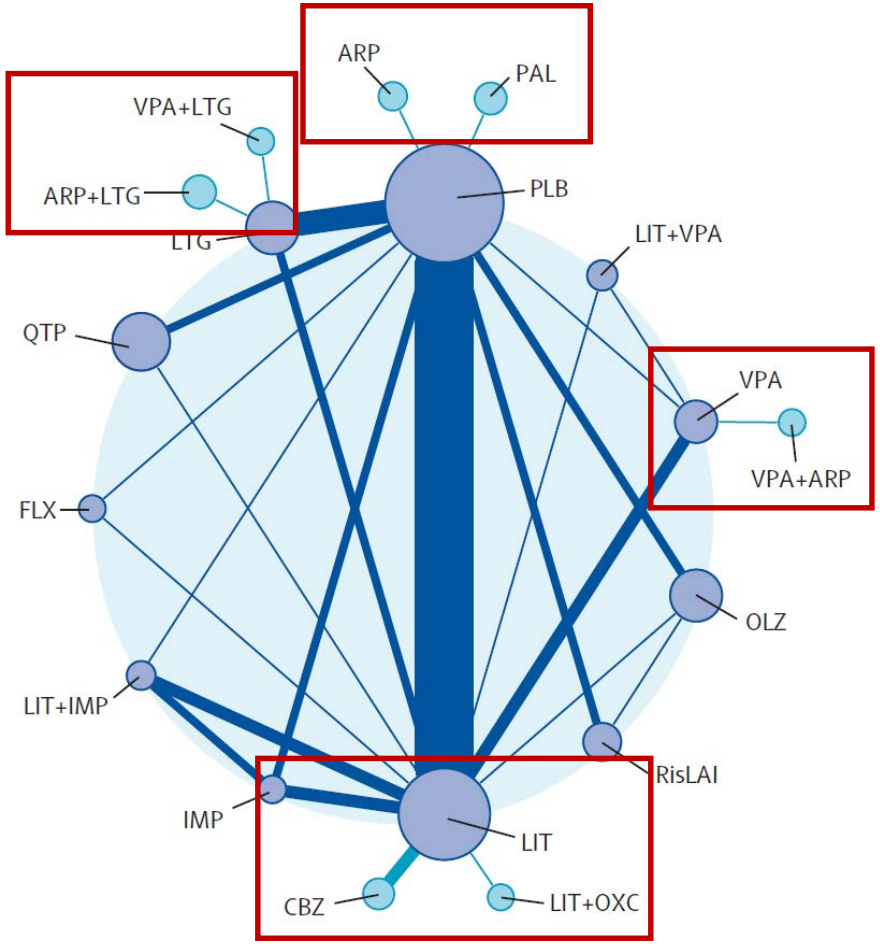
Figure 2: **Network of eligible comparisons for the multiple-treatment meta-analysis for efficacy (response rate)**
The width of the lines is proportional to the number of trials comparing each pair of treatments, and the size of each node is proportional to the number of randomised participants (sample size). The network of eligible comparisons for acceptability (dropout rate) analysis is similar.

- エビデンスのネットワークの構造を視覚的に表すグラフ
- ノードの大きさは、個々の治療群のサンプルサイズの総数に比例
- 線の太さは、個々の対比較（直接比較）のサンプルサイズの総数に比例
- それぞれの治療群・対比較のパスが、全体の結果にどの程度、寄与するかを理解するために有用なグラフ

Cipriani, Furukawa, Salanti et al. (2009)

営利目的でのご利用や無断転載はご遠慮ください

Network Plot: Another Example



- 双極性障害の維持治療における薬物療法のネットワークメタアナリシス
- 17治療（N=6846）を比較した33のランダム化比較試験を統合
- Triangle Loopを形成しない孤立したノードは、間接比較のエビデンスを共有できないため、直接比較のエビデンス以上に情報量の増加はなく、比較的不安定な推定結果が得られやすいので注意！！

Miura, Noma, Furukawa et al. (2014)

営利目的でのご利用や無断転載はご遠慮ください

Statistical Analysis: Definition of Outcomes

- We chose a dichotomous primary outcome mainly for clinical reasons. We used both the number of patients who responded and the number of patients who dropped out to have hard outcome measures of both treatment efficacy and acceptability. We used response rate instead of a continuous symptom score for efficacy analysis to make the interpretation of results easier for clinicians. When dichotomous efficacy outcomes were not reported, but baseline scores, endpoint means, and standard deviations (SD) of the depression rating scales (such as HDRS or MADRS) were provided, we estimated the number of patients responding to treatment at 8 weeks (range 6–12 weeks) with a validated imputation method.

Cipriani, Furukawa, Salanti et al. (2009)

営利目的でのご利用や無断転載はご遠慮ください

Statistical Analysis: Preliminary Assessments

- First, we did pair-wise meta-analyses by synthesizing studies that compared the same interventions with a random-effects model to incorporate the assumption that different studies assessed different, yet related, treatment effects. We used visual inspection of the forest plots to investigate the possibility of statistical heterogeneity, and the I^2 statistic. We did the analyses using Stata version 9.

Cipriani, Furukawa, Salanti et al. (2009)

営利目的でのご利用や無断転載はご遠慮ください

Statistical Analysis: Network Meta-Analysis

- Second, we did a random-effects model within a Bayesian framework using Markov chain Monte Carlo methods in WinBUGS (MRC Biostatistics Unit, Cambridge, UK). We modelled the binary outcomes in every treatment group of every study, and specified the relations among the odds ratios (ORs) across studies making different comparisons. This method combines direct and indirect evidence for any given pair of treatments.

Cipriani, Furukawa, Salanti et al. (2009)

営利目的でのご利用や無断転載はご遠慮ください

参考：ベイズ統計学 Bayesian Statistics

- Bayesの定理
  - $\Pr(H|D) \propto \Pr(D|H)\Pr(H)$
- 関心のある治療効果のパラメータへの事前情報を事前分布として取り込み、データの情報と併せて推測を行う統計学の流派（⇔ 頻度論 Frequentist）
- 複雑な構造を持つモデルに対しても、Markov Chain Monte Carlo法などの計算技法を用いて、統一的な枠組みで分析を行うことができる

Thomas Bayes (1701-1761)

営利目的でのご利用や無断転載はご遠慮ください

NMAの利点①：推定精度・検出力の向上

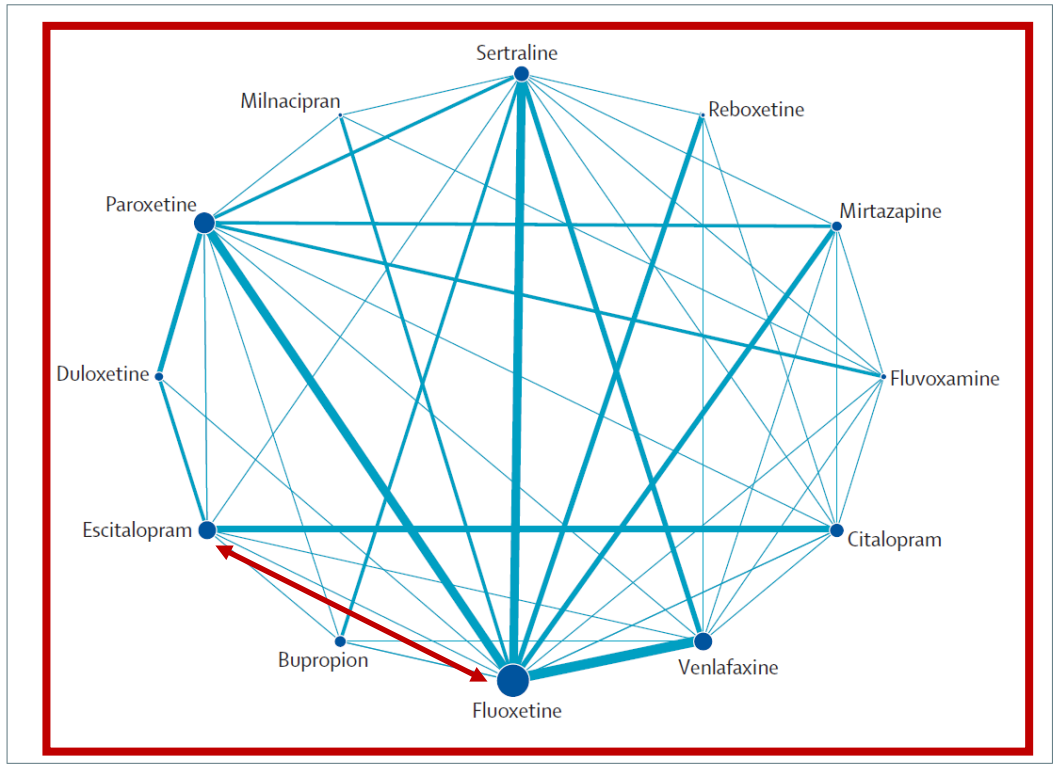


Figure 2: **Network of eligible comparisons for the multiple-treatment meta-analysis for efficacy (response rate)**
The width of the lines is proportional to the number of trials comparing each pair of treatments, and the size of each node is proportional to the number of randomised participants (sample size). The network of eligible comparisons for acceptability (dropout rate) analysis is similar.

Cipriani, Furukawa, Salanti et al. (2009)

- 間接比較のエビデンスを加えることにより、より精度の高い評価が可能になる
- Escitalopram vs Fluoxetine
  - 直接比較　2試験
    OR 1.23, 95%CI: 0.87, 1.75
  - 間接比較のエビデンスを加えると…
    OR 1.30, 95%CI: 1.12, 1.50
- 統計学的な情報量が大きくなるため、精度が高まり、有意な結果が得られる

営利目的でのご利用や無断転載はご遠慮ください

NMAの利点②：直接比較のエビデンスがなくても

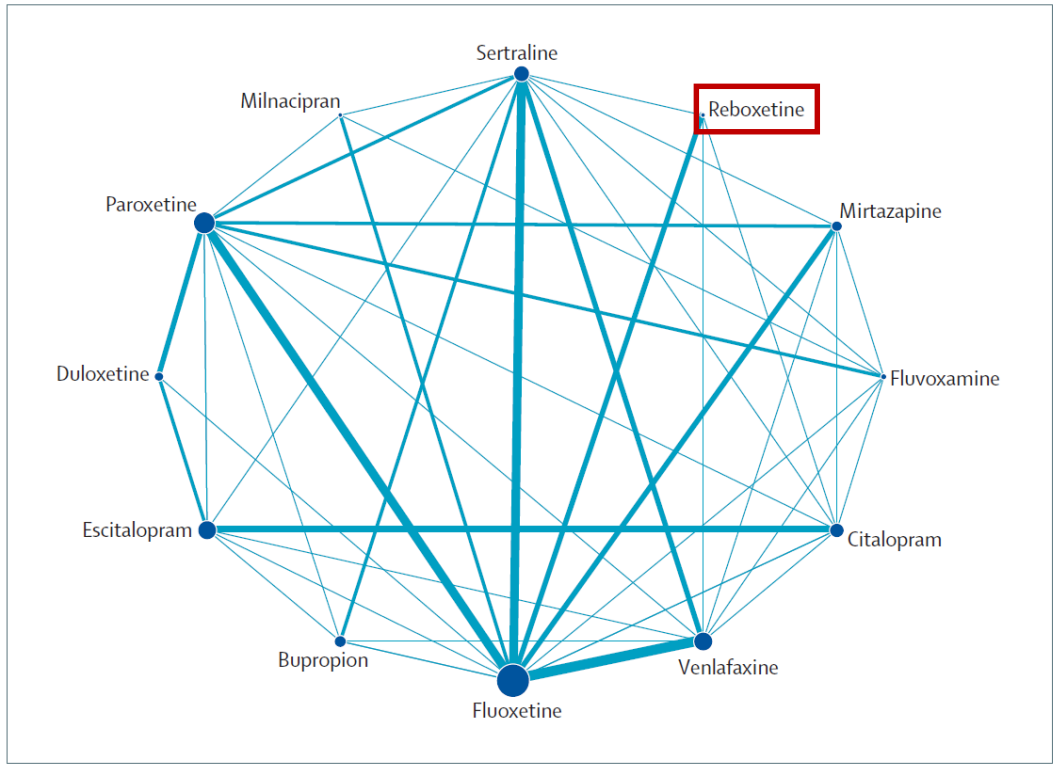


Figure 2: **Network of eligible comparisons for the multiple-treatment meta-analysis for efficacy (response rate)**
The width of the lines is proportional to the number of trials comparing each pair of treatments, and the size of each node is proportional to the number of randomised participants (sample size). The network of eligible comparisons for acceptability (dropout rate) analysis is similar.

Cipriani, Furukawa, Salanti et al. (2009)

- 直接比較をした試験が１つもないパスでも、間接比較のエビデンスによって、ネットワーク上のその他のすべての治療と有効性を比較することができる
- Reboxetineは、ネットワーク上で、直接比較のパスは、Sertraline, Fluoxetine, Venlafaxine, Citalopramの4薬剤しかないが…

営利目的でのご利用や無断転載はご遠慮ください

Reboxetine vs.

	OR (95%CI)
Bupropion	0.63 (0.46, 0.83)
Citalopram	0.61 (0.47, 0.80)
Duloxetine	0.68 (0.50, 0.95)
Escitalopram	0.51 (0.39, 0.68)
Fluoxetine	0.68 (0.53, 0.86)
Fluvoxamine	0.69 (0.50, 0.97)
Milnacipran	0.67 (0.46, 0.97)
Mirtazapine	0.49 (0.36, 0.66)
Paroxetine	0.67 (0.50, 0.86)
Sertraline	0.54 (0.41, 0.71)

他の11薬剤すべてに有意に劣るという結果に

営利目的でのご利用や無断転載はご遠慮ください

NMAの利点③：Comparative Effectivenessの評価

- ネットワーク上のすべての治療を比較した治療効果の差の推定値，信頼区間，P値を得ることができる（オッズ比，ハザード比など）
- 治療の順位付け（ランキング）を行い、対象となった治療間でのそれぞれの順位を推定することができる

営利目的でのご利用や無断転載はご遠慮ください

対比較オッズ比の推定値，信頼区間

Efficacy (response rate) (95% CI) | Comparison | Acceptability (dropout rate) (95% CI)

BUP	1·00 (0·78–1·28)	0·75 (0·55–1·01)	1·06 (0·86–1·32)	0·89 (0·74–1·08)	0·73 (0·53–1·00)	0·87 (0·58–1·24)	0·87 (0·66–1·14)	0·81 (0·65–1·00)	0·62 (0·45–0·86)	1·01 (0·82–1·27)	0·84 (0·68–1·02)
0·98 (0·78–1·23)	CIT	0·75 (0·55–1·02)	1·07 (0·86–1·31)	0·90 (0·73–1·09)	0·73 (0·54–0·99)	0·87 (0·60–1·24)	0·87 (0·66–1·15)	0·81 (0·65–1·01)	0·62 (0·45–0·84)	1·02 (0·81–1·28)	0·84 (0·67–1·06)
1·09 (0·83–1·43)	1·12 (0·87–1·44)	DUL	1·43 (1·09–1·85)	1·19 (0·91–1·57)	0·98 (0·67–1·41)	1·16 (0·77–1·73)	1·16 (0·83–1·61)	1·08 (0·84–1·40)	0·83 (0·57–1·22)	1·36 (1·01–1·83)	1·12 (0·84–1·50)
0·82 (0·67–1·01)	0·84 (0·70–1·01)	0·75 (0·60–0·93)	ESC	0·84 (0·70–1·01)	0·69 (0·50–0·94)	0·81 (0·55–1·15)	0·81 (0·62–1·07)	0·76 (0·62–0·93)	0·58 (0·43–0·81)	0·95 (0·77–1·19)	0·78 (0·64–0·97)
1·08 (0·90–1·29)	1·10 (0·93–1·31)	0·99 (0·79–1·24)	1·32 (1·12–1·55)	FLU	0·82 (0·62–1·07)	0·97 (0·69–1·32)	0·97 (0·77–1·21)	0·91 (0·79–1·05)	0·70 (0·53–0·92)	1·14 (0·96–1·36)	0·94 (0·81–1·09)
1·10 (0·83–1·47)	1·13 (0·86–1·47)	1·01 (0·74–1·38)	1·35 (1·02–1·76)	1·02 (0·81–1·30)	FVX	1·18 (0·76–1·75)	1·18 (0·87–1·61)	1·10 (0·84–1·47)	0·85 (0·57–1·26)	1·38 (1·03–1·89)	1·14 (0·86–1·54)
1·07 (0·77–1·48)	1·09 (0·78–1·50)	0·97 (0·69–1·38)	1·30 (0·95–1·78)	0·99 (0·74–1·31)	0·97 (0·68–1·37)	MIL	0·99 (0·69–1·53)	0·94 (0·68–1·31)	0·72 (0·48–1·10)	1·17 (0·84–1·72)	0·97 (0·69–1·40)
0·79 (0·72–1·00)	0·80 (0·63–1·01)	0·72 (0·54–0·94)	0·96 (0·76–1·19)	0·73 (0·60–0·88)	0·71 (0·55–0·92)	0·74 (0·53–1·01)	MIR	0·93 (0·75–1·17)	0·72 (0·51–1·03)	1·17 (0·91–1·51)	0·97 (0·76–1·23)
1·06 (0·87–1·30)	1·08 (0·90–1·30)	0·97 (0·78–1·20)	1·30 (1·10–1·53)	0·98 (0·86–1·12)	0·96 (0·76–1·23)	1·00 (0·74–1·33)	1·35 (1·11–1·64)	PAR	0·77 (0·56–1·05)	1·25 (1·04–1·52)	1·03 (0·86–1·24)
1·60 (1·20–2·16)	1·63 (1·25–2·14)	1·46 (1·05–2·02)	1·95 (1·47–2·59)	1·48 (1·16–1·90)	1·45 (1·03–2·02)	1·50 (1·03–2·18)	2·03 (1·52–2·78)	1·50 (1·16–1·98)	REB	1·63 (1·19–2·24)	1·34 (0·99–1·83)
0·87 (0·72–1·05)	0·88 (0·72–1·07)	0·79 (0·62–1·01)	1·06 (0·88–1·27)	0·80 (0·69–0·93)	0·79 (0·61–1·01)	0·81 (0·60–1·11)	1·10 (0·90–1·36)	0·82 (0·69–0·96)	0·54 (0·41–0·71)	SER	0·82 (0·67–1·00)
0·85 (0·70–1·01)	0·86 (0·71–1·05)	0·77 (0·60–0·99)	1·03 (0·86–1·24)	0·78 (0·68–0·90)	0·77 (0·59–0·99)	0·79 (0·58–1·08)	1·08 (0·87–1·33)	0·79 (0·67–0·94)	0·53 (0·40–0·69)	0·98 (0·82–1·16)	VEN

Cipriani, Furukawa, Salanti et al. (2009)

営利目的でのご利用や無断転載はご遠慮ください

Fluoxetineを基準とした結果

	Efficacy (response rate) OR (95% CI)	Acceptability (dropout rate) OR (95% CI)
Bupropion	0·93 (0·77–1·11)	1·12 (0·92–1·36)
Citalopram	0·91 (0·76–1·08)	1·11 (0·91–1·37)
Duloxetine	1·01 (0·81–1·27)	0·84 (0·64–1·10)
Escitalopram	0·76 (0·65–0·89)*	1·19 (0·99–1·44)
Fluvoxamine	1·02 (0·81–1·30)	0·82 (0·62–1·07)
Milnacipran	0·99 (0·74–1·31)	0·97 (0·69–1·32)
Mirtazapine	0·73 (0·60–0·88)*	0·97 (0·77–1·21)
Paroxetine	0·98 (0·86–1·12)	0·91 (0·79–1·05)
Reboxetine	1·48 (1·16–1·90)*	0·70 (0·53–0·92)*
Sertraline	0·80 (0·69–0·93)*	1·14 (0·96–1·36)
Venlafaxine	0·78 (0·68–0·90)*	0·94 (0·81–1·09)

OR=odds ratio. CI=credibilty interval. *p<0·05. For efficacy, OR higher than 1 favours fluoxetine. For acceptability, OR lower than 1 favours fluoxetine.

Table 4: Efficacy and acceptability using fluoxetine as reference compound

営利目的でのご利用や無断転載はご遠慮ください

治療効果の順位の推定結果

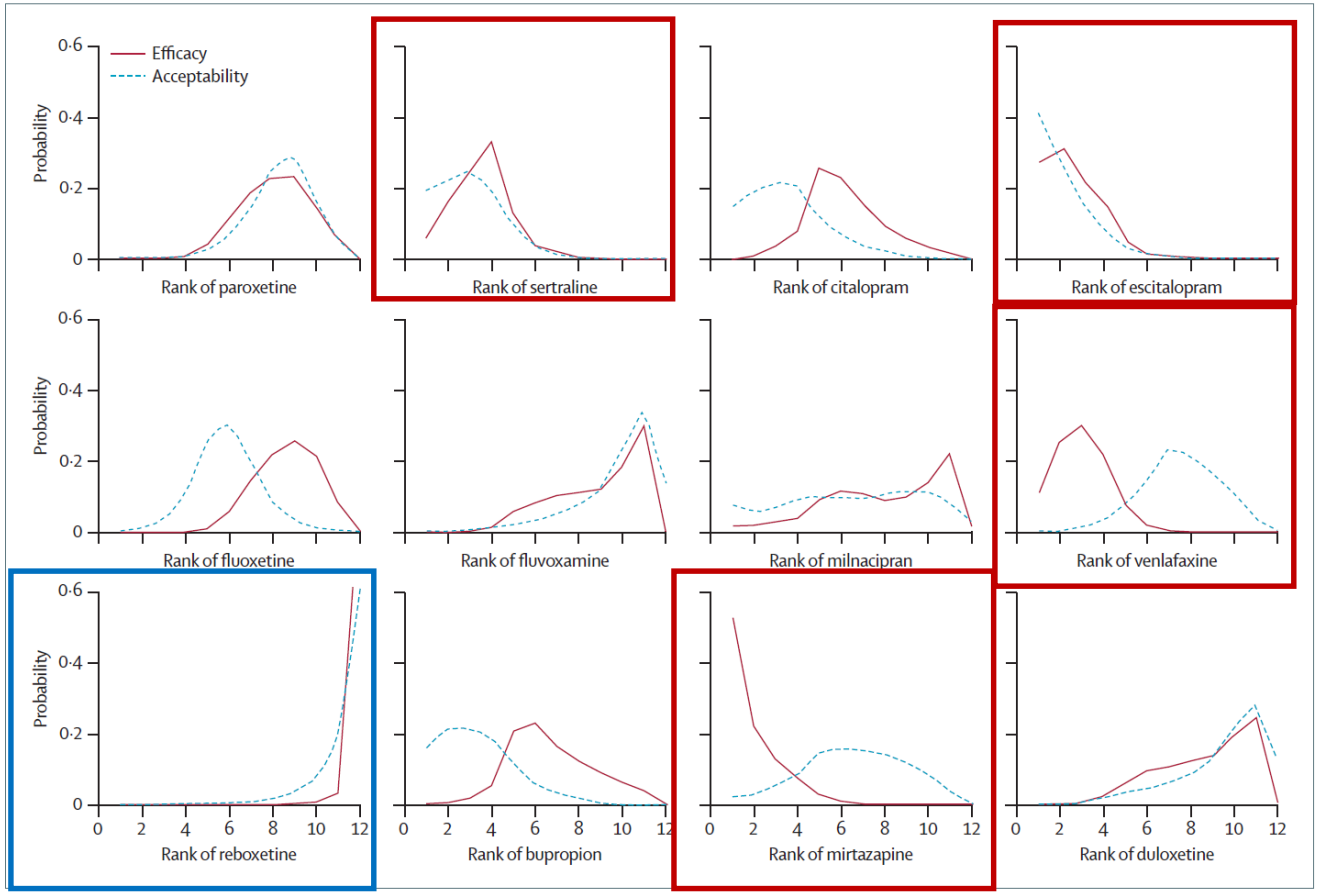


Figure 4: **Ranking for efficacy (solid line) and acceptability (dotted line)**
Ranking indicates the probability to be the best treatment, the second best, the third best, and so on, among the 12 antidepressants.

Cipriani, Furukawa, Salanti et al. (2009)

営利目的でのご利用や無断転載はご遠慮ください

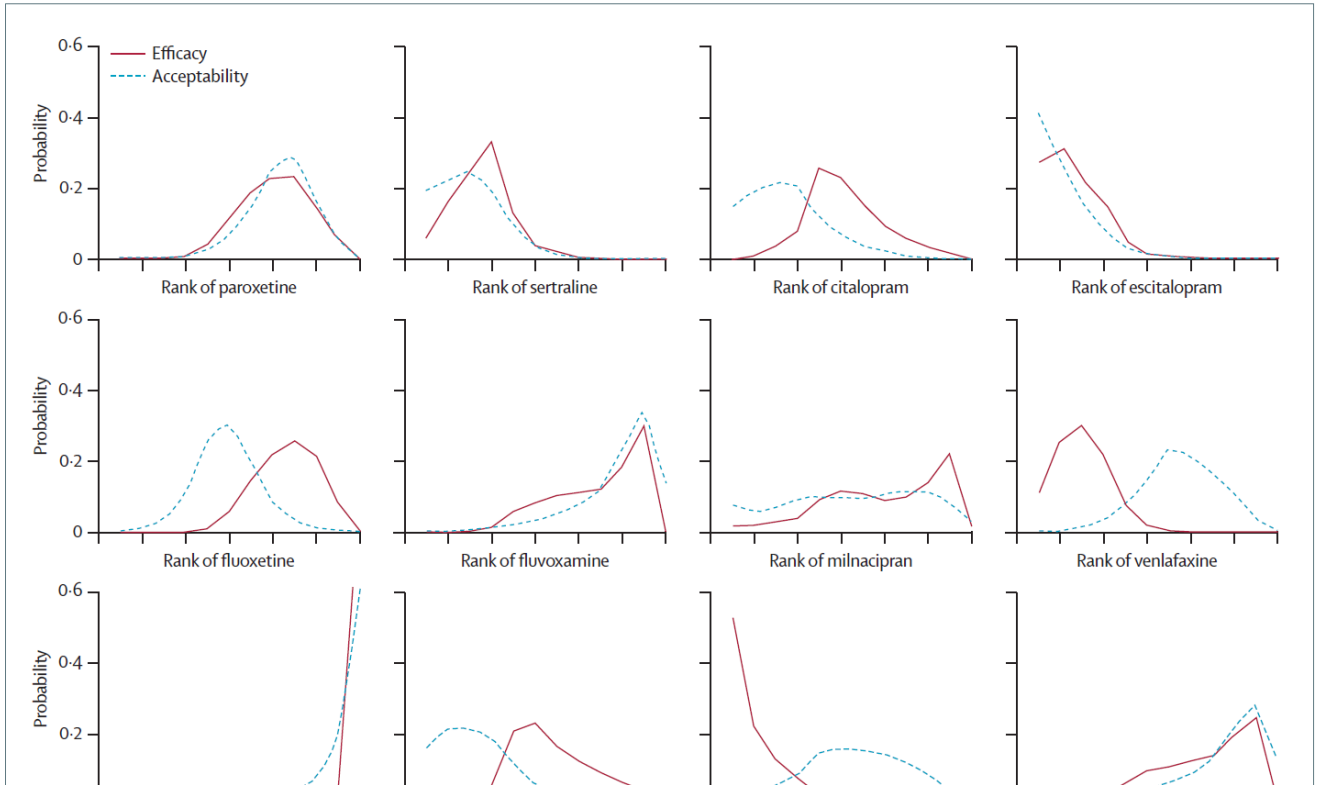
Comparative Efficacyについての結果

- Figure 3 summarises the results of the multiple treatments meta-analysis. Escitalopram, mirtazapine, sertraline, and venlafaxine were significantly more efficacious than duloxetine, fluoxetine, fluvoxamine, paroxetine, and reboxetine (even though less clear benefits were noted with sertraline than with escitalopram, venlafaxine, and mirtazapine when comparing with duloxetine and fluvoxamine, with the CI for OR slightly more than 1).
- Reboxetine was significantly less efficacious than all the other 11 antidepressants.
- These findings arise from 66 simultaneous comparisons and about three statistically significant findings might be expected by chance alone.

Cipriani, Furukawa, Salanti et al. (2009)

営利目的でのご利用や無断転載はご遠慮ください

治療効果の順位の推定結果



治療間の順位はわかりやすく魅力的な指標だが、僅差のものの微小な差の変化で覆る不安定な指標なので、解釈には注意が必要である！！

Cipriani, Furukawa, Salanti et al. (2009)

営利目的でのご利用や無断転載はご遠慮ください

Fluoxetineを基準とした結果

	Efficacy (response rate) OR (95% CI)	Acceptability (dropout rate) OR (95% CI)
Bupropion	0·93 (0·77–1·11)	1·12 (0·92–1·36)
Citalopram	0·91 (0·76–1·08)	1·11 (0·91–1·37)
Duloxetine	1·01 (0·81–1·27)	0·84 (0·64–1·10)
Escitalopram	0·76 (0·65–0·89)*	1·19 (0·99–1·44)
Fluvoxamine	1·02 (0·81–1·30)	0·82 (0·62–1·07)
Milnacipran	0·99 (0·74–1·31)	0·97 (0·69–1·32)
Mirtazapine	0·73 (0·60–0·88)*	0·97 (0·77–1·21)
Paroxetine	0·98 (0·86–1·12)	0·91 (0·79–1·05)
Reboxetine	1·48 (1·16–1·90)*	0·70 (0·53–0·92)*
Sertraline	0·80 (0·69–0·93)*	1·14 (0·96–1·36)
Venlafaxine	0·78 (0·68–0·90)*	0·94 (0·81–1·09)

OR=odds ratio. CI=credibilty interval. *p<0·05. For efficacy, OR higher than 1 favours fluoxetine. For acceptability, OR lower than 1 favours fluoxetine.

Table 4: Efficacy and acceptability using fluoxetine as reference compound

営利目的でのご利用や無断転載はご遠慮ください

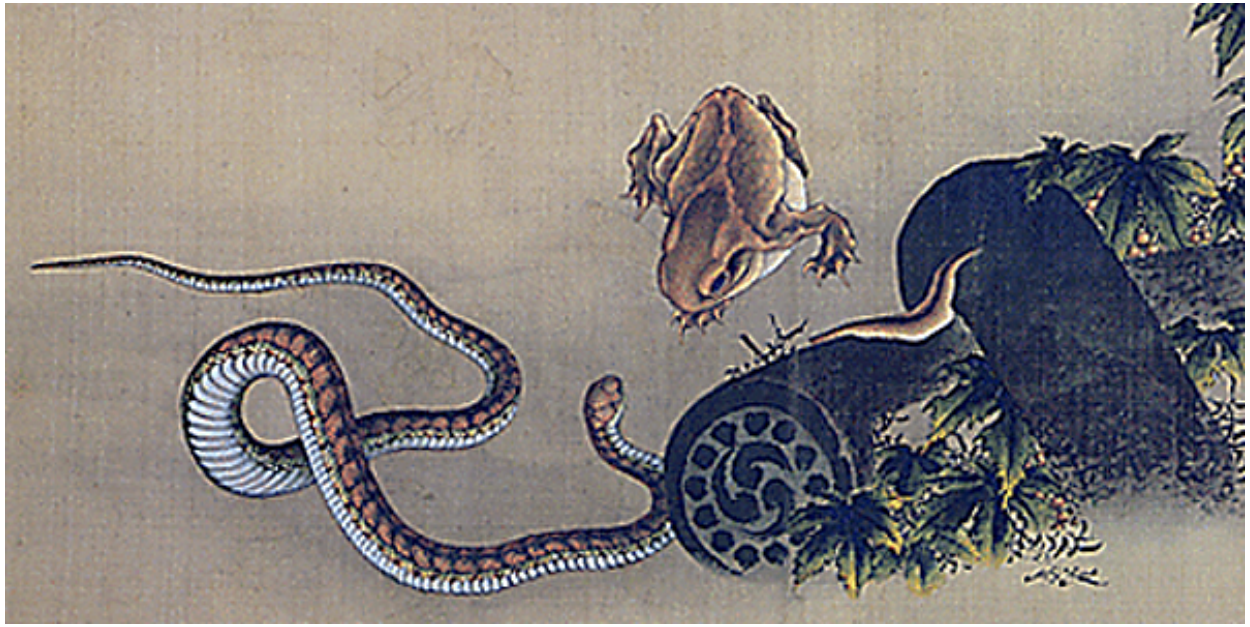
【警告】これより話がさらに難しくなります

- 月曜は社畜に強く
- 社畜は無職に強く
- 無職は月曜に強い

https://twitter.com/LoversCats/status/732412420010840064 より一部改変

営利目的でのご利用や無断転載はご遠慮ください

うまい話には罠があるとは言いますが…

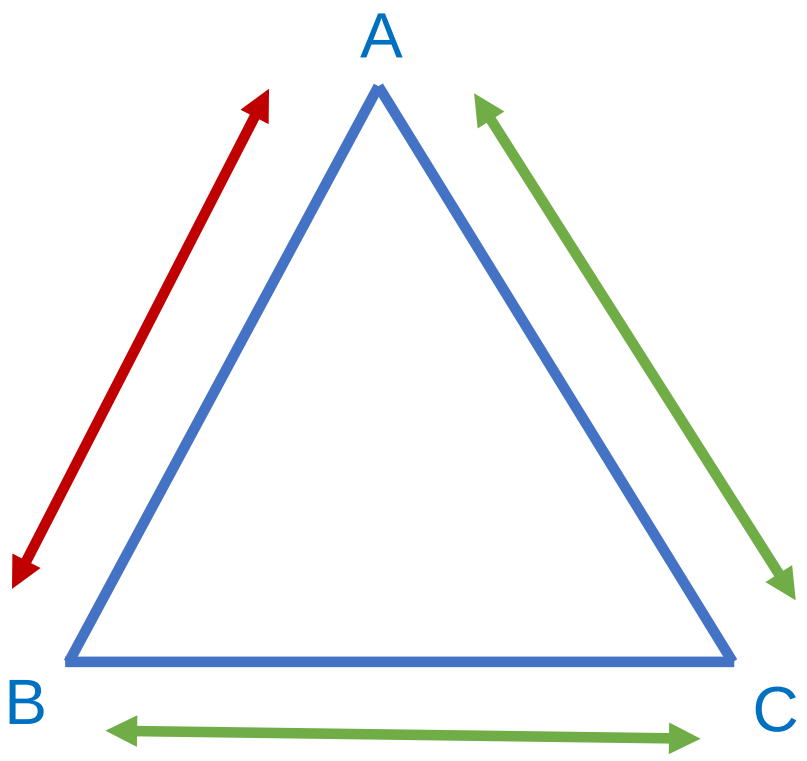


- 蝦蛆食蛇、蛇食蛙、蛙食蝦蛆、互相食也

《「関尹子」三極から》蛇はなめくじをおそれ、なめくじは蛙（かえる）をおそれ、蛙は蛇をおそれること。転じて、三者が互いに牽制し合って、それぞれが自由に動けない状態。

営利目的でのご利用や無断転載はご遠慮ください

ネットワーク上の比較の妥当性



- NMAにおいて、直接比較・間接比較のエビデンスを統合して、治療間の妥当な比較を行うためには、直接比較・間接比較のパスにおける治療間の差が一致していなくてはならない
  - 直接比較：A vs B
  - 間接比較：B vs C, C vs A
- 2つのパスにおけるA-B間の差が一致しなくては、NMAにおける治療の比較の妥当性は失われてしまう！！

Salanti (2012), Dias et al. (2013)

営利目的でのご利用や無断転載はご遠慮ください

Inconsistency

- Consistency: ネットワーク上のすべてのパスにおいて、治療間のEffect Sizeの差が一致（consistent）すること
  - ネットワーク上の任意のペアA-Bを結ぶすべてのパスにおけるEffect Sizeの差が互いに一致して整合すること
  - より広義の概念として、Transitivityと言われることも
- Inconsistency: Lack of consistency
  - ネットワーク上の治療の比較の妥当性が、前提として成り立たないということに！！

営利目的でのご利用や無断転載はご遠慮ください

Local inconsistency on the network

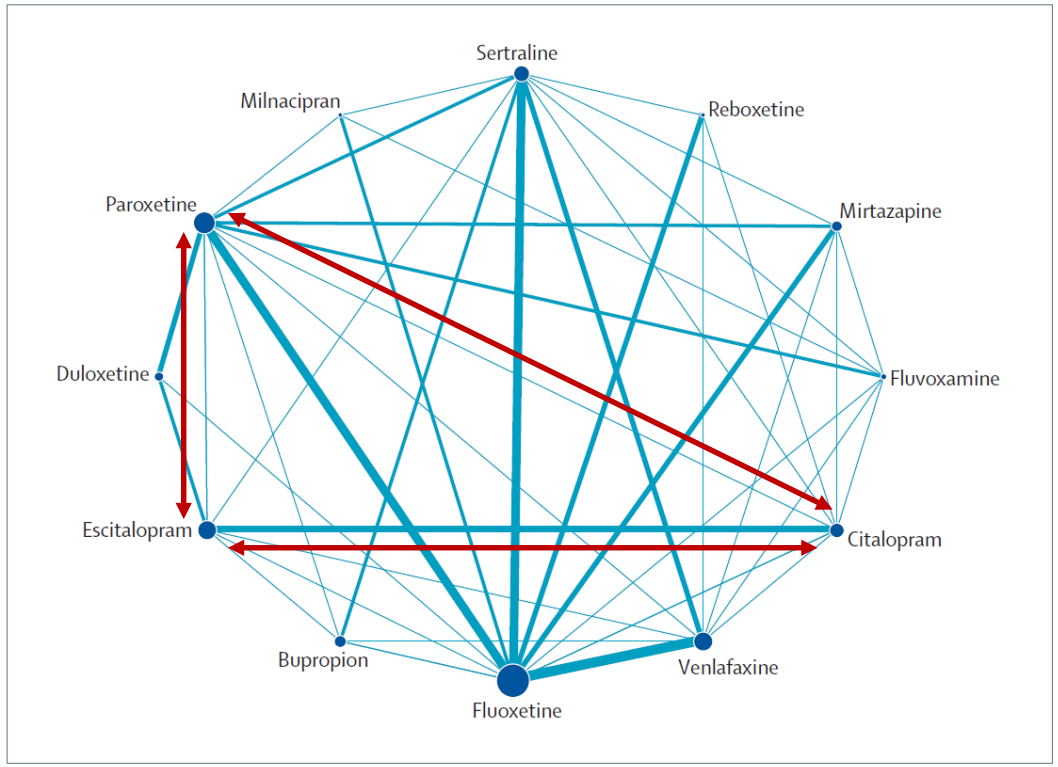


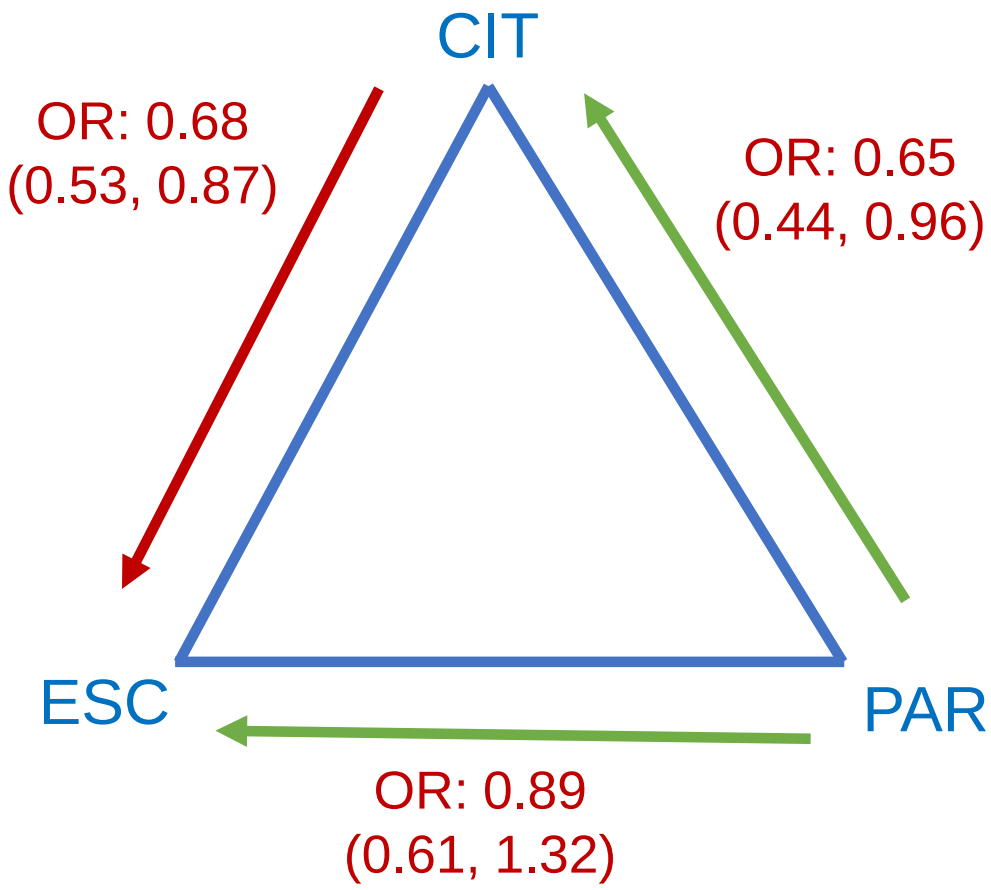
Figure 2: **Network of eligible comparisons for the multiple-treatment meta-analysis for efficacy (response rate)**
The width of the lines is proportional to the number of trials comparing each pair of treatments, and the size of each node is proportional to the number of randomised participants (sample size). The network of eligible comparisons for acceptability (dropout rate) analysis is similar.

Cipriani, Furukawa, Salanti et al. (2009)

- ネットワーク上の任意のTriangleに対して、３辺のパスにおけるEffect SizeはConsistentか？
- NMAの解析モデルは、ネットワーク上のすべてのパスにConsistencyを仮定しているので、1つでもそれが崩れると、妥当性が成り立たない
- Triangle上のInconsistencyは、簡単な手順で検定をすることができる
- 例えば、CIT-ESC-PARのTriangleでは…

営利目的でのご利用や無断転載はご遠慮ください

CIT-ESC-PARのTriangle Loop



- CIT vs. ESCの直接比較では、ESCのほうが有意に反応率が高い
- CIT vs. PARでは、CITのほうが有意に反応率が高い
- ESCは、PARよりも有意に反応率が高い…
- わけではなく、ESCとPARの直接比較では、有意差が出ず、ORも 0.89 ほどしかない
- いわゆる「三竦」に近い状態！！
- ESC-PAR-CIT間のパス上のEffect Sizeは一致しない？？

Cipriani et al. (2009)

営利目的でのご利用や無断転載はご遠慮ください

Local Inconsistencyの検定

- Triangle上のループの直接比較のパスと間接比較のパスのEffect Sizeが一致するかどうかは、検定で評価することができる

$$Z = \frac{\hat{\delta}_{Direct} - \hat{\delta}_{Indirect}}{\sqrt{V(\hat{\delta}_{Direct}) + V(\hat{\delta}_{Indirect})}} \sim N(0,1)$$

- 有意になったTriangle Loopは、Inconsistencyが疑われるので、バイアスを生じさせる要因がないかどうかを詳細に検討する必要がある

Bucher et al. (1999), Chaimani et al. (2013)

営利目的でのご利用や無断転載はご遠慮ください

Results: Inconsistencyの評価

- Analysis indicated statistical incoherence in three out of 70 comparisons of direct with indirect evidence for response rate (**paroxetine-citalopram-escitalopram; fluvoxamine-venlafaxine-mirtazapine; and sertraline-fluoxetine-bupropion**)....
- ネットワーク上で、直接比較の存在するTriangle Loopsは全部で70あり、そのすべてで、Inconsistencyを評価する検定を行った
- 先ほどのPAR-CIT-ESCのループを含む、3つのループが有意になった
- Inconsistencyの原因となる要因についての検討が行われている

Cipriani et al. (2009)

営利目的でのご利用や無断転載はご遠慮ください

Results: Inconsistencyの評価

- Data extraction and data entering were correct, and we could not identify any important variable that differed across comparisons in those loops; however, the number of included studies was small.
- These numbers are compatible with chance because about six significant findings would be expected out of 133 statistical tests.
- 詳細な検討が行われたが、重要な要因は同定できなかったと結論
- Efficacy, Acceptabilityのアウトカムを含めて、6/133の検定が有意になった（Efficacy Outcomeについては、3/70=4.2%）
- 一方、仮に、すべてのLoopsがConsistentであったとしても、5%水準の検定では、平均的に5%（70個中3~4個の検定）は、誤って有意になる
- ランダムな誤差の範囲に含まれる、と結論されている

Cipriani et al. (2009)

営利目的でのご利用や無断転載はご遠慮ください

Side-splitting法

- それでも、ネットワーク上の対比較ごとに、直接・間接エビデンスが一致するかどうかは関心の対象である
  - Local Inconsistency Testと同様、ネットワーク上のどのパスにInconsistencyの原因があるかを調べるためにも、重要
- ネットワーク上の対比較のパス（e.g., A vs. B）ごとに、治療効果の差の推定値を、直接・間接エビデンスに分離して、要約する方法が Side-splitting法である

Dias et al. (2010), White (2015), Noma et al. (2017)

営利目的でのご利用や無断転載はご遠慮ください

ネットワーク全体の間接エビデンスの情報

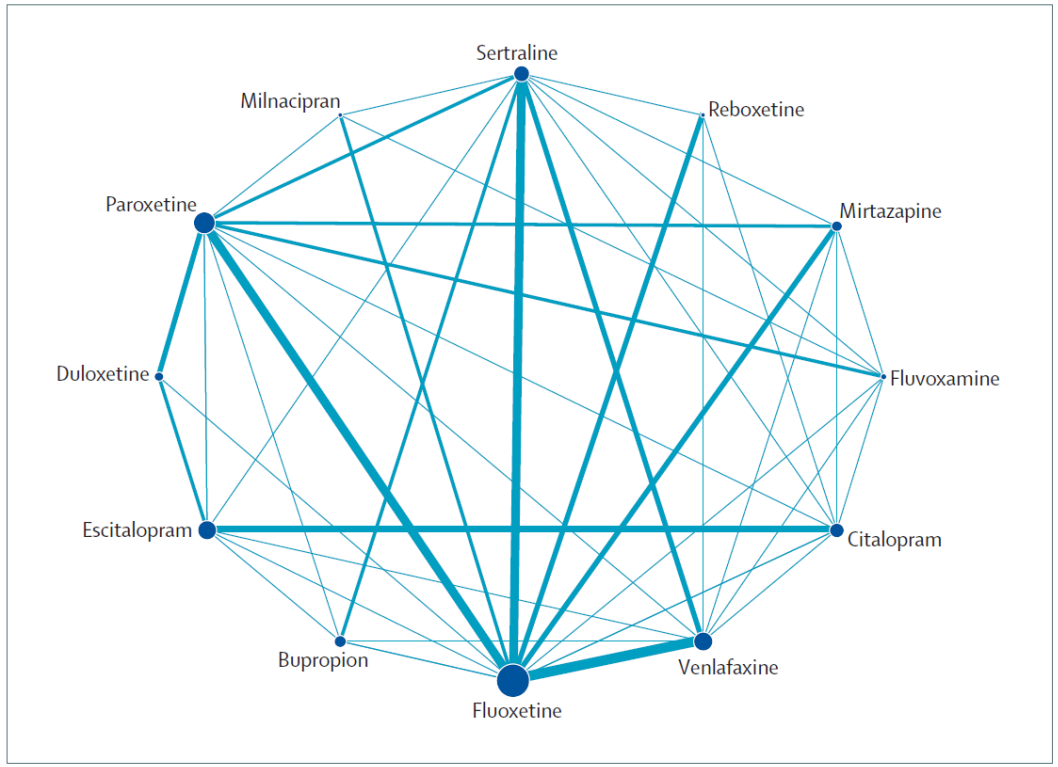


Figure 2: **Network of eligible comparisons for the multiple-treatment meta-analysis for efficacy (response rate)**
The width of the lines is proportional to the number of trials comparing each pair of treatments, and the size of each node is proportional to the number of randomised participants (sample size). The network of eligible comparisons for acceptability (dropout rate) analysis is similar.

Cipriani, Furukawa, Salanti et al. (2009)

- 局所的なTriangle Loopsに限定するのではなく、ネットワーク全体での間接比較のエビデンスをすべてまとめて、要約したオッズ比の推定値にまとめることはできないか？
- Noma et al. (2017) では、この間接エビデンスを要約する方法と、Inconsistencyを評価するための検定を開発した

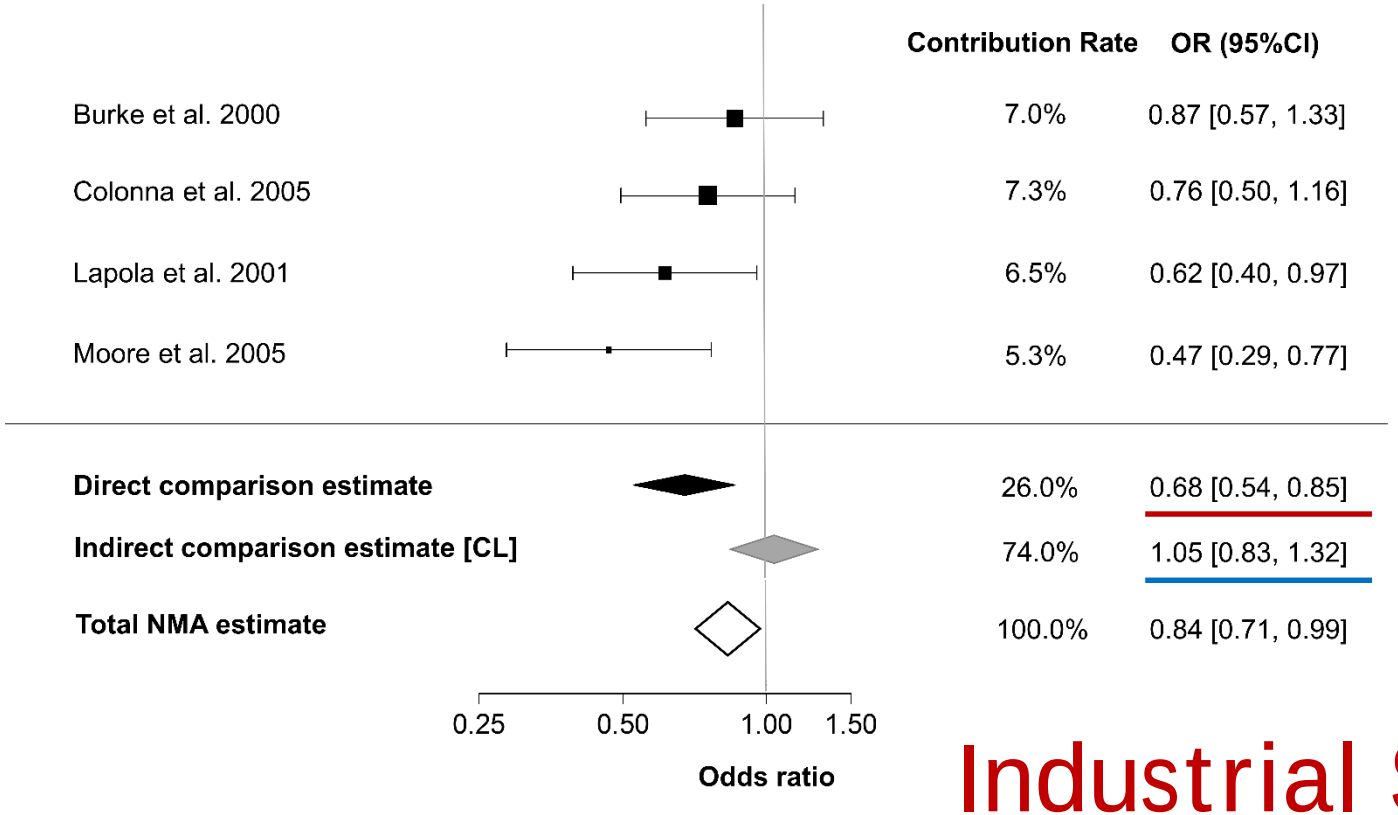
営利目的でのご利用や無断転載はご遠慮ください

Selected Results (by the Order of P-values)

	Number of direct comparison	OR for the whole network	Contribution rate of direct comparison	OR for direct comparison	OR for indirect comparison	P-value
Citalopram vs Escitalopram	4	0.84 (0.71, 0.99)	26.0%	0.68 (0.54, 0.85)	1.05 (0.83, 1.32)	0.009
Mirtazapine vs Venlafaxine	2	1.07 (0.88, 1.29)	8.3%	1.53 (1.03, 2.27)	0.95 (0.77, 1.19)	0.041
Bupropion vs Fluoxetine	3	1.06 (0.91, 1.25)	8.0%	0.82 (0.61, 1.11)	1.19 (0.98, 1.44)	0.042
Citalopram vs Paroxetine	1	1.09 (0.92, 1.29)	6.3%	1.54 (1.03, 2.31)	1.02 (0.85, 1.23)	0.068
Bupropion vs Sertraline	3	0.87 (0.73, 1.03)	16.3%	1.07 (0.79, 1.46)	0.78 (0.63, 0.97)	0.103
Fluoxetine vs Sertraline	8	0.81 (0.71, 0.94)	29.3%	0.70 (0.56, 0.89)	0.89 (0.74, 1.06)	0.123
Fluvoxamine vs Milnacipran	1	0.98 (0.71, 1.36)	11.4%	0.57 (0.26, 1.24)	1.10 (0.77, 1.57)	0.131

営利目的でのご利用や無断転載はご遠慮ください

Escitalopram vs. Citalopram



P=0.009

Industrial Sponsorship Bias??

Noma et al. (2017)

営利目的でのご利用や無断転載はご遠慮ください

まとめ

- 近年のComparative Effectiveness Researchの振興からも、ネットワークメタアナリシスは、システマティックレビューのスタンダードな方法のひとつとして、今後、ますます普及していくことが予想される
- 方法論そのものは魅力的であるが、強い理論的仮定に依存する方法であり、さまざまなバイアスが含まれるリスクがある
- PRISMA-NMA extended version（Hutton et al., 2015）も公表されており、NMA特有の問題に配慮して、結果の報告を行う必要がある

営利目的でのご利用や無断転載はご遠慮ください

文献

- Bucher, H. C., Guyatt, G. H., Griffith, L. E., and Walter, S. D. (1997). The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *Journal of Clinical Epidemiology* **50**, 683-691.
- Caldwell, D. M., Ades, A. E., and Higgins, J. P. (2005). Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *British Medical Journal* **331**, 897-900.
- Chaimani, A., Higgins, J. P., Mavridis, D., Spyridonos, P., and Salanti, G. (2013). Graphical tools for network meta-analysis in STATA. *PLoS One* **8**, e76654.
- Cipriani, A., Furukawa, T. A., Salanti, G., *et al.* (2009). Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* **373**, 746-758.
- Dias, S., Welton, N. J., Caldwell, D. M., and Ades, A. E. (2010). Checking consistency in mixed treatment comparison meta-analysis. *Statistics in Medicine* **29**, 932-944.
- Dias, S., Welton, N. J., Sutton, A. J., Caldwell, D. M., Lu, G., and Ades, A. E. (2013). Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. *Medical Decision Making* **33**, 641-656.

営利目的でのご利用や無断転載はご遠慮ください

- Hutton, B., Salanti, G., Caldwell, D. M., *et al.* (2015). The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Annals of Internal Medicine* **162**, 777-784.
- Li, T., and Becker, L. (2013). Comparing multiple intervention; Workshop 1. In 21st Cochrane Colloquium. Quebec, Canada, September, 2013.
- Mitka, M. (2010). US government kicks off program for comparative effectiveness research. *Journal of the American Medical Association* **304**, 2230-2231.
- Miura, T., Noma, H., Furukawa, T. A., *et al.* (2014). Comparative efficacy and tolerability of pharmacological treatments in the maintenance treatment of bipolar disorder: a systematic review and network meta-analysis. *Lancet Psychiatry* **1**, 351-359.
- Montgomery, S., Hansen, T., and Kasper, S. (2011). Efficacy of escitalopram compared to citalopram: a meta-analysis. *International Journal of Neuropsychopharmacology* **14**, 261-268.
- Murad, M. H., Drake, M. T., Mullan, R. J., *et al.* (2012). Clinical review. Comparative effectiveness of drug treatments to prevent fragility fractures: a systematic review and network meta-analysis. *Journal of Clinical Endocrinology & Metabolism* **97**, 1871-1880.

営利目的でのご利用や無断転載はご遠慮ください

- Noma, H., Tanaka, S., Matsui, S., Cipriani, A., and Furukawa, T. A. (2017). Quantifying indirect evidence in network meta-analysis. *Statistics in Medicine* **36**, 917-927.
- Salanti, G. (2012). Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Research Synthesis Methods* **3**, 80-97.
- White, I. R. (2015). Network meta-analysis. *The Stata Journal* **15**, 951-985.

営利目的でのご利用や無断転載はご遠慮ください