ネットワークメタアナリシスにおける Inconsistencyと種々のバイアス

野間　久史
統計数理研究所 データ科学研究系，リスク解析戦略研究センター
第3回データサイエンスラウンドテーブル会議
平成28年2月29日
e-mail: noma@ism.ac.jp
URL: http://www.ism.ac.jp/~noma/

12種類の新世代抗うつ薬のNMA

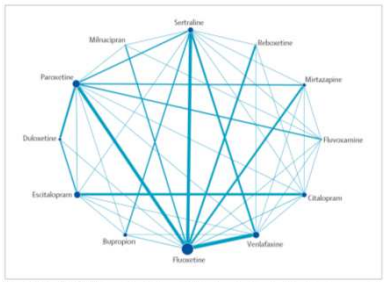


Figure 2: Network of eligible comparisons for the multiple-treatment meta-analysis for efficacy (response rate)
The width of the lines is proportional to the number of trials comparing each pair of treatments, and the size of each node is proportional to the number of randomised participants (sample 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種類の抗うつ薬を比較したランダム化臨床試験（117試験；25,298 人）を対象
- 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）
- ベイズ流の階層モデルを用いた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を比較・評価

Cipriani, Furukawa, Salanti et al. (2009)

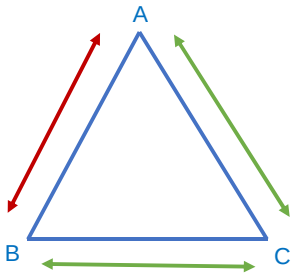
対比較オッズ比の推定値，95%信用区間

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 (0·67–1·44)	1·16 (0·77–1·72)	1·16 (0·83–1·61)	1·08 (0·84–1·40)	0·83 (0·57–1·22)	1·36 (1·03–1·83)	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 (0·77–1·15)	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 (0·69–1·37)	0·97 (0·77–1·23)	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 (0·82–1·61)	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 (0·74–1·33)	0·97 (0·68–1·37)	MIL	0·99 (0·69–1·43)	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 (0·82–1·38)	1·08 (0·90–1·30)	0·97 (0·78–1·20)	1·30 (1·19–1·53)	0·98 (0·86–1·12)	0·96 (0·76–1·23)	1·00 (0·74–1·33)	1·35 (1·13–1·64)	PAR	0·77 (0·58–1·05)	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 (0·65–0·93)	0·79 (0·61–1·01)	0·81 (0·60–1·11)	1·10 (0·90–1·36)	0·82 (0·65–0·96)	0·54 (0·43–0·71)	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)

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)

直接・間接エビデンスの Inconsistency

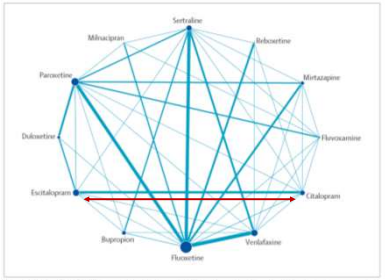
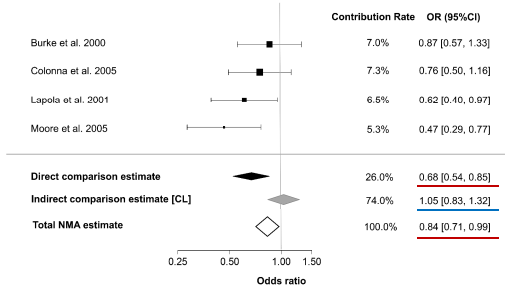


Figure 2: Network of eligible comparisons for the multiple-treatment meta-analysis for efficacy (response rate)
The width of the lines is proportional to the number of trials comparing each pair of treatments, and the size of each node is proportional to the number of randomised participants (sample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

- ネットワーク全体での直接比較のエビデンス・間接比較のエビデンスのInconsistencyを評価することができる
- Escitalopram-Citalopramのパスで、両者のInconsistencyを評価してみると…

Cipriani et al. (2009), Noma et al. (2015)

Escitalopram vs. Citalopram



P=0.009

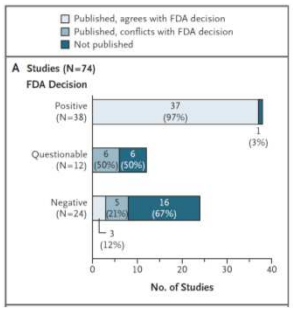
Cipriani et al. (2009), Noma et al. (2015)

Post-hocな考察

- Escitalopramは、Citalopramと同一の企業により開発された、後発の薬剤であり、直接比較の試験は、この企業がスポンサーになっている
  - 直接比較のエビデンスでは、後発のEscitalopramの優越性が明確に示されている
- 間接比較（ネットワーク上のそれ以外の比較試験の情報）のエビデンスでは、全体として、ESC, CITの差は、Moderateなものに（点推定値としては、ほとんど差はない）

※ もちろん、By Chanceでこのような結果が得られることもある。

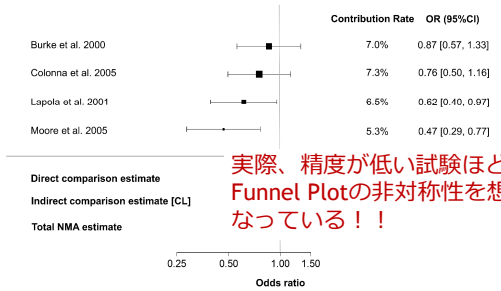
Inconsistencyの原因は？？



Turner et al. (2008)

- “Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy” by Turner et al., NEJM, 2008; 358:252-260.
- FDAの臨床試験レジストレーションに登録されていた抗うつ薬の臨床試験の出版状況が系統的に評価されている
- Positiveな結果が得られていた試験は、多くのものが論文化され、出版に至っていたものの、そうでない試験は、出版されていないものが多かった
- NMAでは、このような直接比較試験の出版バイアスも、ネットワーク上のInconsistencyを生じさせ、全体の解析結果にバイアスを生じさせる可能性がある！！

Escitalopram vs. Citalopram



実際、精度が低い試験ほど、推定値が大きく、Funnel Plotの非対称性を想起させる分布になっている！！

Cipriani et al. (2009), Noma et al. (2016)


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 LAPERRIE, 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 the outcome) or negative (indicating that the outcome with treatment was the same or worse than without it) at the conventional level of statistical significance (P<0.05).

Results We identified 12 large randomized, controlled trials and 19 meta-analyses addressing the same questions. For a total of 40 primary and secondary outcomes, agreement between the meta-analyses and the large clinical trials was only fair (kappa = 0.35; 95 percent confidence interval, 0.06 to 0.64). The positive predictive value of the meta-analyses was 68 percent, and the negative predictive value 67 per-

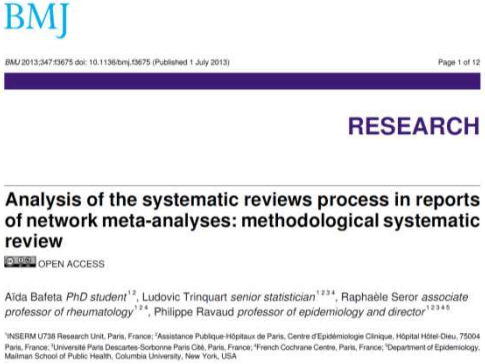
LARGE randomized, controlled trials are generally considered the gold standard in evaluations of the efficacy of clinical interventions. However, since such trials are not always available, clinicians increasingly rely on meta-analysis to support their choice of clinical strategies. Critics have emphasized the intrinsic weaknesses of meta-analysis.[1-3] 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.[4] reviewed 30 meta-analyses of various interventions in perinatal medicine from the Cochrane data base. They recalculated the results of each meta-analysis after removing the largest trial from the analysis and then compared the results with those of the large trial that had been removed. They found a kappa of 0.46 to 0.53 and a positive predictive value of 50 to 67 percent. We compared the results of a series of systematically compiled large randomized, controlled trials with those of the relevant meta-analyses that had been published previously.

- （NMAでない）従来のメタアナリシスでも、さまざまなバイアスによって、結果の再現性が得られないことが
- BMJ, Lancet, JAMA, NEJMなどの一流誌の中だけでも、早期に行われたメタアナリシスと、後続の大規模なランダム化臨床試験の結果が食い違ったという例がかなりの数に上るという報告が行われたことも！！
- ネットワークメタアナリシスでは、通常のメタアナリシス以上に、強い理論的仮定と多種多様なバイアスが結果に影響を及ぼすことに！！

LeLorier et al. (1997)

Practices on Network Meta-Analysis


BMJ

BMJ 2013;347:f3675 doi: 10.1136/bmj.f3675 (Published 1 July 2013) Page 1 of 12

RESEARCH

Analysis of the systematic reviews process in reports of network meta-analyses: methodological systematic review

OPEN ACCESS

Aida Bafeta *PhD student*[1 2], Ludovic Trinquart *senior statistician*[1 2 3 4], Raphaèle Seror *associate professor of rheumatology*[1 2 4], Philippe Ravaud *professor of epidemiology and director*[1 2 3 4 5]

[1]INSERM U738 Research Unit, Paris, France; [2]Assistance Publique-Hôpitaux de Paris, Centre d'Epidémiologie Clinique, Hôpital Hôtel-Dieu, 75004 Paris, France; [3]Université Paris Descartes-Sorbonne Paris Cité, Paris, France; [4]French Cochrane Centre, Paris, France; [5]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, USA

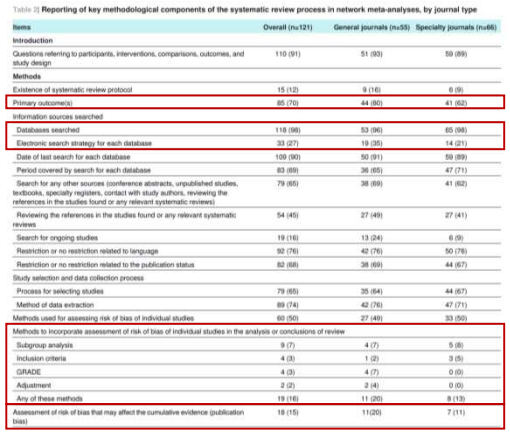


Table 2| **Reporting of key methodological components of the systematic review process in network meta-analyses, by journal type**

Items	Overall (n=121)	General journals (n=55)	Specialty journals (n=66)
Introduction			
Questions referring to participants, interventions, comparisons, outcomes, and study design	110 (91)	51 (93)	59 (89)
Methods			
Existence of systematic review protocol	15 (12)	9 (16)	6 (9)
Primary outcome(s)	85 (70)	44 (80)	41 (62)
Information sources searched			
Databases searched	118 (98)	53 (96)	65 (98)
Electronic search strategy for each database	33 (27)	19 (35)	14 (21)
Date of last search for each database	109 (90)	50 (91)	59 (89)
Period covered by search for each database	83 (69)	36 (65)	47 (71)
Search for any other sources (conference abstracts, unpublished studies, textbooks, specialty registers, contact with study authors, reviewing the references in the studies found or any relevant systematic reviews)	79 (65)	38 (69)	41 (62)
Reviewing the references in the studies found or any relevant systematic reviews	54 (45)	27 (49)	27 (41)
Search for ongoing studies	19 (16)	13 (24)	6 (9)
Restriction or no restriction related to language	92 (76)	42 (76)	50 (76)
Restriction or no restriction related to the publication status	82 (68)	38 (69)	44 (67)
Study selection and data collection process			
Process for selecting studies	79 (65)	35 (64)	44 (67)
Method of data extraction	89 (74)	42 (76)	47 (71)
Methods used for assessing risk of bias of individual studies	60 (50)	27 (49)	33 (50)
Methods to incorporate assessment of risk of bias of individual studies in the analysis or conclusions of review			
Subgroup analysis	9 (7)	4 (7)	5 (8)
Inclusion criteria	4 (3)	1 (2)	3 (5)
GRADE	4 (3)	4 (7)	0 (0)
Adjustment	2 (2)	2 (4)	0 (0)
Any of these methods	19 (16)	11 (20)	8 (13)
Assessment of risk of bias that may affect the cumulative evidence (publication bias)	18 (15)	11(20)	7 (11)

Befeta et al. (2013)

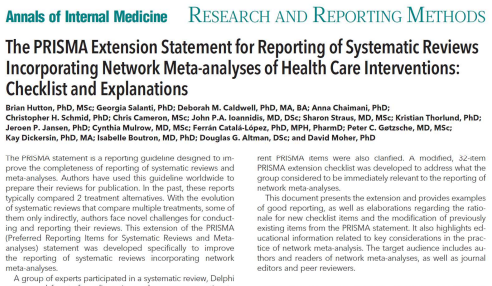
Table 3| **Inadequate quality of conduct of the systematic review process with adequate reporting in network meta-analyses, by journal type**

Item (reported and of inadequate quality of conduct)	Overall	General journal	Specialty journal
Electronic search of only one bibliographic database	13/118 (11)	8/53 (15)	5/65 (8)
Restriction of study selection based on publication status	16/82 (20)	6/38 (16)	10/44 (23)
Lack of independent duplicate study selection	34/79 (43)	13/37 (35)	21/42 (50)
Lack of independent duplicate data extraction	19/89 (21)	10/41 (24)	9/48 (19)

Data are number (%)/total number of reports. Denominators of fractions indicate the total number of reports in which the corresponding item was reported.

Befeta et al. (2013)

PRISMA Extended Version for NMA



Annals of Internal Medicine RESEARCH AND REPORTING METHODS

The PRISMA Extension Statement for Reporting of Systematic Reviews Incorporating Network Meta-analyses of Health Care Interventions: Checklist and Explanations

Brian Hutton, PhD, MSc; Georgia Salanti, PhD; Deborah M. Caldwell, PhD, MA, BA; Anna Chaimani, PhD; Christopher H. Schmid, PhD; Chris Cameron, MSc; John P.A. Ioannidis, MD, DSc; Sharon Straus, MD, MSc; Kristian Thorlund, PhD; Jeroen P. Jansen, PhD; Cynthia Mulrow, MD, MSc; Ferrán Catalá-López, PhD, MPH, PharmD; Peter C. Gøtzsche, MD, MSc; Kay Dickersin, PhD, MA; Isabelle Boutron, MD, PhD; Douglas G. Altman, DSc; and David Moher, PhD

The PRISMA statement is a reporting guideline designed to improve the completeness of reporting of systematic reviews and meta-analyses. Authors have used this guideline worldwide to prepare their reviews for publication. In the past, these reports typically compared 2 treatment alternatives. With the evolution of systematic reviews that compare multiple treatments, some of them only indirectly, authors face novel challenges for conducting and reporting their reviews. This extension of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) statement was developed specifically to improve the reporting of systematic reviews incorporating network meta-analyses.

A group of experts participated in a systematic review, Delphi

rent PRISMA items were also clarified. A modified, 32-item PRISMA extension checklist was developed to address what the group considered to be immediately relevant to the reporting of network meta-analyses.

This document presents the extension and provides examples of good reporting, as well as elaborations regarding the rationale for new checklist items and the modification of previously existing items from the PRISMA statement. It also highlights educational information related to key considerations in the practice of network meta-analysis. The target audience includes authors and readers of network meta-analyses, as well as journal editors and peer reviewers.

個々のNMAの論文を批評的に読み解く際の参考にも

Hutton et al. (2015)

STATAのNMA解析モジュール

ネットワークメタアナリシスの解析モジュール

野間 久史
統計数理研究所
e-mail: noma[at]ism.ac.jp
URL: http://www.ism.ac.jp/~noma/
平成 28 年 2 月 11 日

ネットワークメタアナリシスは，ベイズ流の枠組みのもとで，マルコフ連鎖モンテカルロ法（Markov Chain Monte Carlo）などの複雑な計算技法を利用した解析が一般的であったが，近年では，多変量メタアナリシスの枠組みのもとでの頻度論的な解析方法も徐々に普及しつつある．特に，STATA では，Ian White 博士による network モジュールを使用することによって，簡単なコマンドでネットワークメタアナリシスの計算が実行できるようになっている．また，Chaimani et al. (2013) で紹介されている種々のグラフィックツールも，network graph パッケージで使用できるようになった．本稿では，

Inconsistencyや種々のバイアスの評価ツールも、
一連のモジュールに含まれています

http://www.ism.ac.jp/~noma/STATA_network.pdf

まとめ

- 解析モジュールの開発・整備も進んでおり、近い将来、統計家でなくても、NMAは実施可能になると思われる
- NMAは、複数の治療が対象となり、解析の妥当性が担保されるためには、強い理論的仮定が必要。さらに、バイアスの要素も多種多様になっていることに注意！！
- バイアスはゼロにすることはできないが、それを最小限に留めるための研究方法を正しく採用すること
- 種々のバイアスのリスクについて、批評的に吟味すること（そのための解析・評価ツールの活用，感度解析の実施）

参考文献

- Bafeta, A., Trinquart, L., Seror, R. and Ravaud, P. (2013). Analysis of the systematic reviews process in reports of network meta-analyses: Methodological systematic review. *BMJ* **347,** f3675.
- 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., Sutton, A. J., *et al*. (2013). Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. *Medical Decision Making* **33**, 641-656.
- 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.
- LeLorier J, Gregoire G, Benhaddad A, *et al*. Discrepancies between meta-analyses and subsequent large randomized, controlled trials. *New England Journal of Medicine* 1997; **337**: 536-542.

- Noma, H., Tanaka, S., Matsui, S., *et al*. (2015). Quantifying indirect evidence in network meta-analysis. East Asia Regional Biometric Conference 2015, Kyushu University, Fukuoka, Japan, December 22, 2016.
- 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.
- Turner EH, Matthews AM, Linardatos E, *et al*. (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine* **358**, 252-260.