

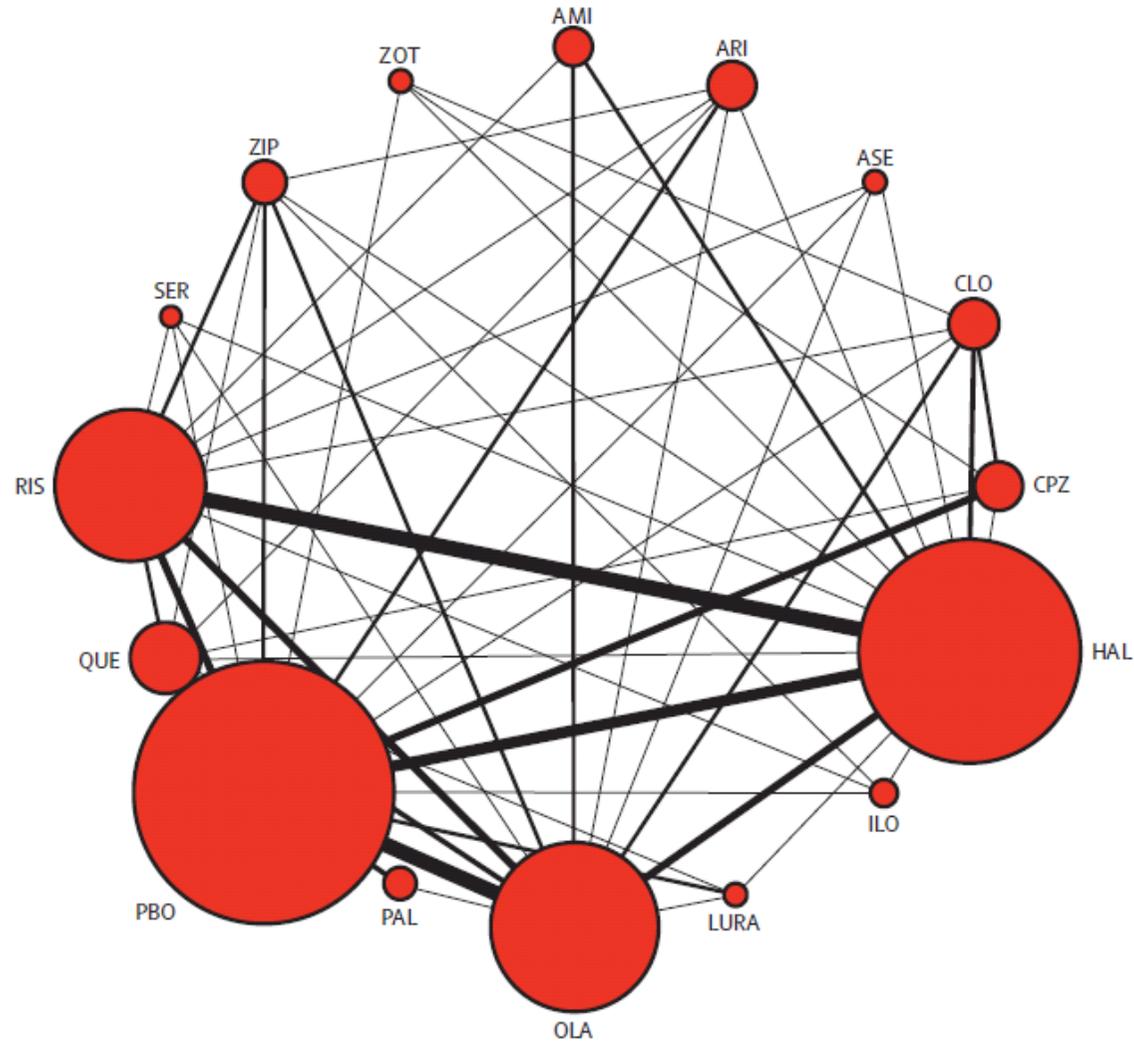
ネットワークメタアナリシス

浅野 淳一(PMDA), 兼清 道雄(エーザイ)
長島 健悟(千葉大学), 藤井 陽介(ファイザー)

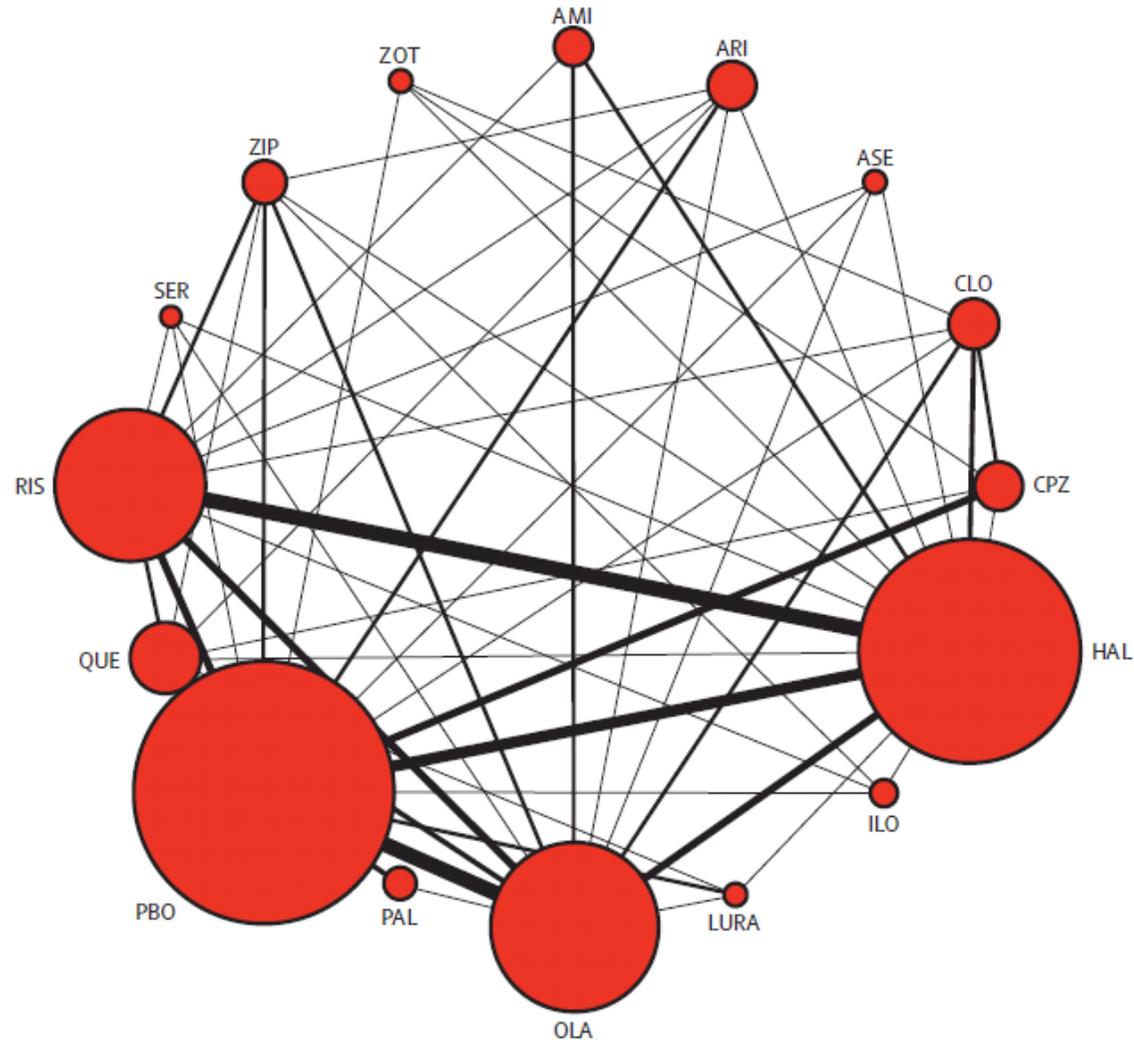
2016年2月29日

第3回 データサイエンスラウンドテーブル

Network Meta-analysis (NMA) は 薬剤比較で結ばれるネットワークでメタアナリシスをする



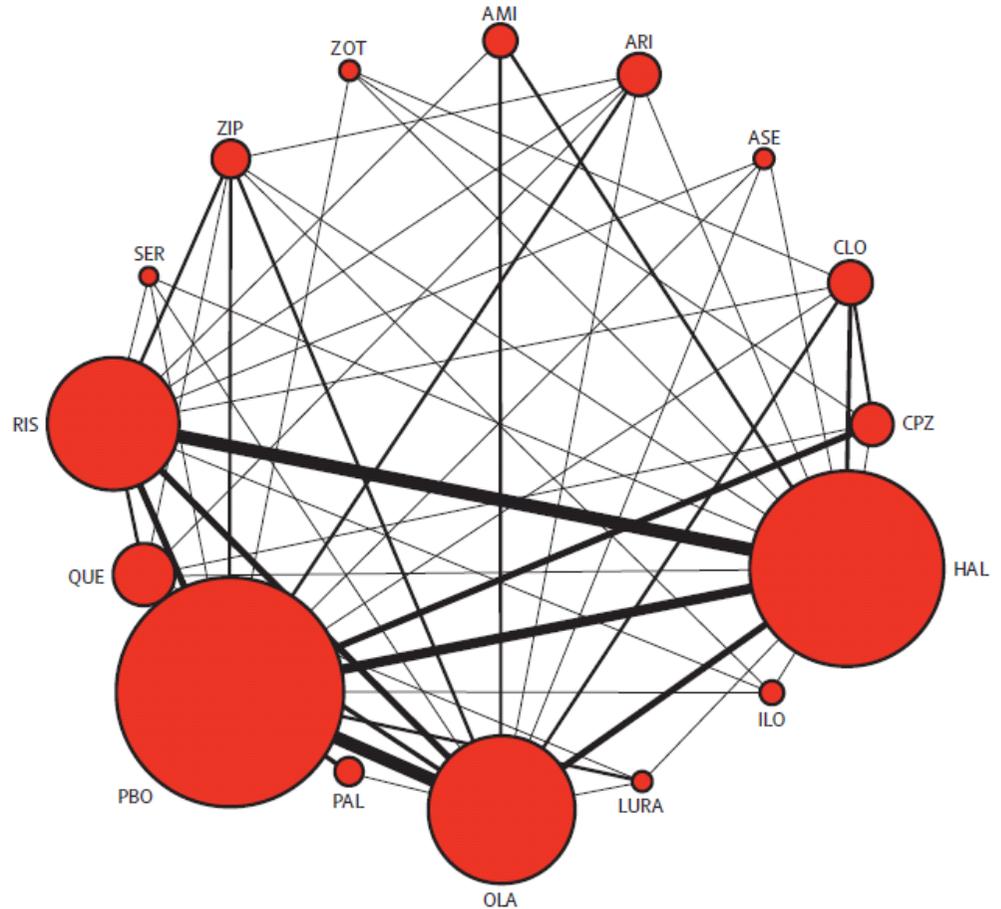
薬剤のどんな組み合わせも比較できる，まるで宝箱？



はたまた、開けてはいけないパンドラの箱？

• デジタル大辞泉より

- ゼウスがパンドラに持たせた、あらゆる災いの詰まった箱(本来は壺)。彼女が地上に着いたとき好奇心から開けたところ、すべての災いが地上に飛び出したが、急いでふたをしたので希望だけが残ったという。



NMAの論文数(Pubmedで検索; 臨床・方法論区別なし)

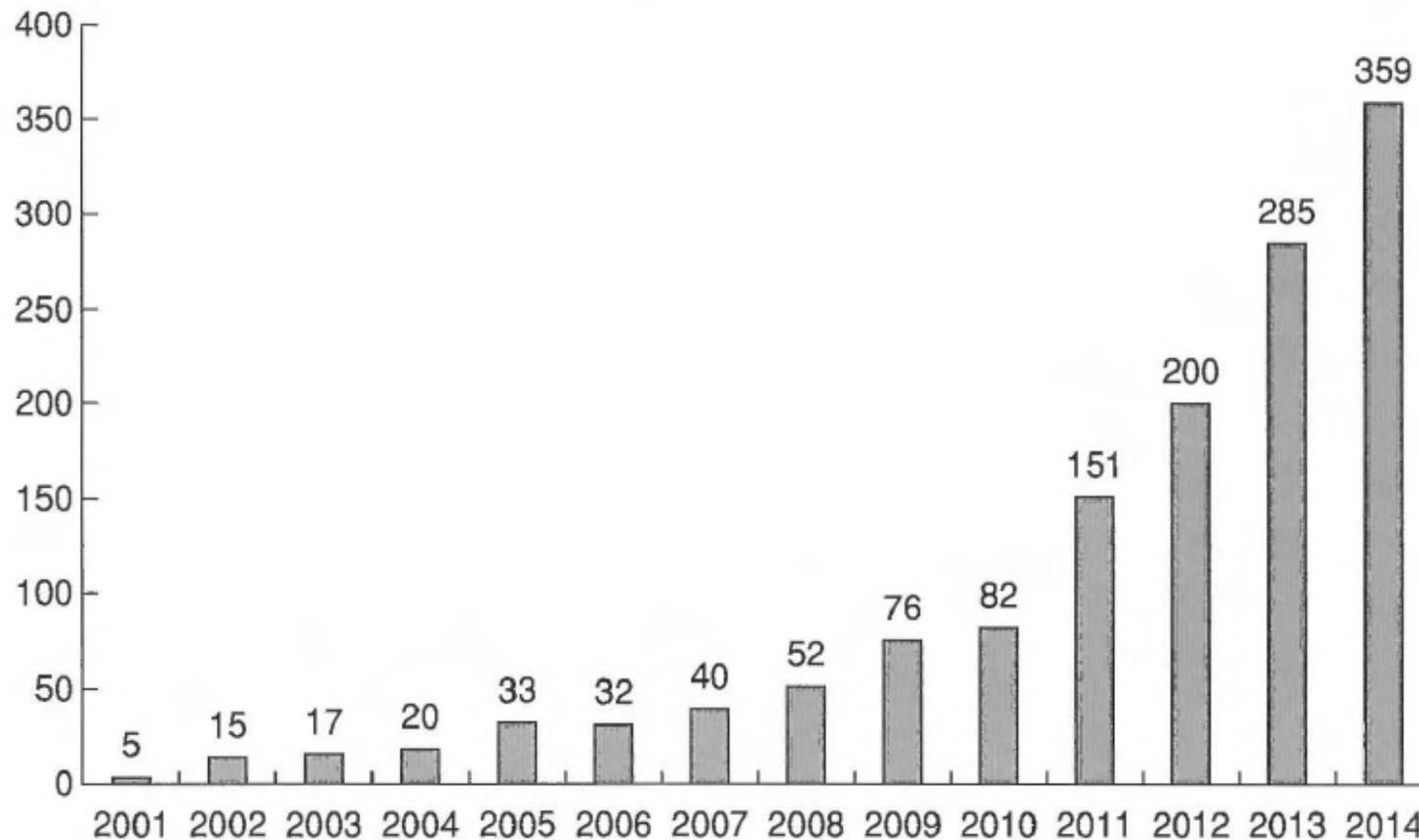


図3 NMAの論文数

Pubmedで検索(検索式: “network meta-analysis” or “multiple treatment comparison meta-analysis” or “mixed comparison meta-analysis” or “multiple comparison meta-analysis” or “mixed treatment comparison meta-analysis”)

[藤井]

On the need for increased rigour and care in the conduct and interpretation of network meta-analyses in drug development

Kevin Carroll^{a*} and Robert Hemmings^b

The rise over recent years in the use of network meta-analyses (NMAs) in clinical research and health economic analysis is little short of meteoric driven, in part, by a desire from decision makers to extend inferences beyond direct comparisons in controlled clinical trials. But is the increased use and reliance of NMAs justified? Do such analyses provide a reliable basis for the relative effectiveness assessment of medicines and, in turn, for critical decisions relating to healthcare access and provisioning? And can such analyses also be used earlier, as part of the evidence base for licensure? Despite several important publications highlighting inherently unverifiable assumptions underpinning NMAs, these assumptions and associated potential for serious bias are often overlooked in the reporting and interpretation of NMAs. A more cautious, and better informed, approach to the use and interpretation of NMAs in clinical research is warranted given the assumptions that sit behind such analyses. Copyright © 2015 John Wiley & Sons, Ltd.

Keywords: network meta-analyses; rigour; bias

本セッションの流れ

1. ディスカッションのための話題提供 (50分)

- 藤井 陽介
- 野間 久史 先生 (統計数理研究所)

休憩 (10分)

2. グループディスカッション1 (60分)

3. グループディスカッション2 (60分)

休憩 (10分)

4. スーパーテーブル(40分)

ディスカッションのための話題提供

藤井 陽介

薬剤の評価，他剤との比較

• 薬剤のライフサイクルにおける薬剤の評価・他剤との比較

– 製薬企業

• 新薬の承認取得のための試験

- 多くの競合薬剤があっても，それらを全て組込む試験は非現実的
- 迅速な承認を狙った開発戦略により，大抵プラセボか標準治療と比較

• 自社の薬剤の位置づけの確認（差別化）

- 承認申請時のエビデンス構築，リスクベネフィット評価[承認，薬価・・・]
- 承認後のエビデンス蓄積[LCM]

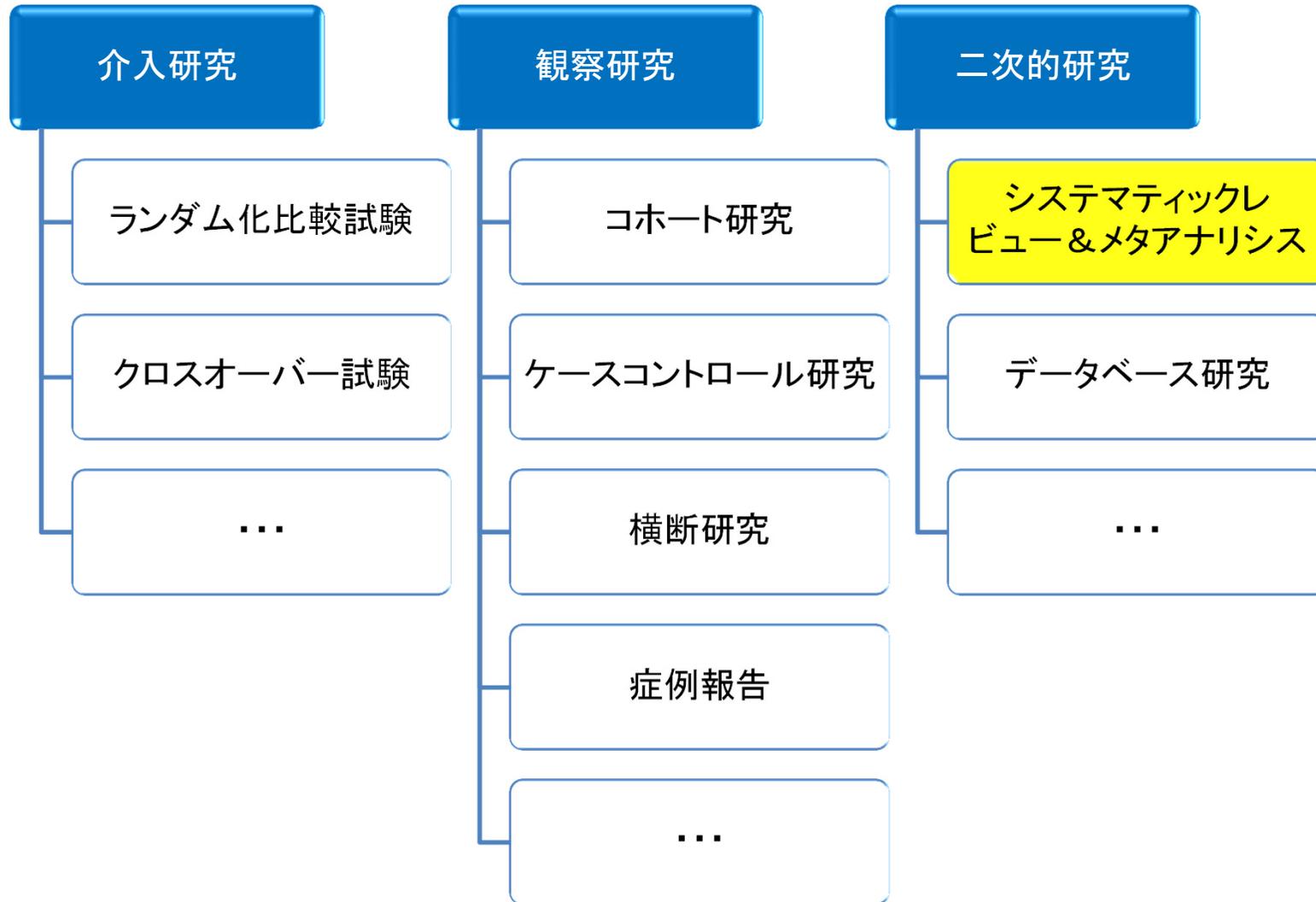
– 医療分野一般

• 意思決定

- 臨床治療ガイドラインの決定
- 保険償還等の医療政策の決定

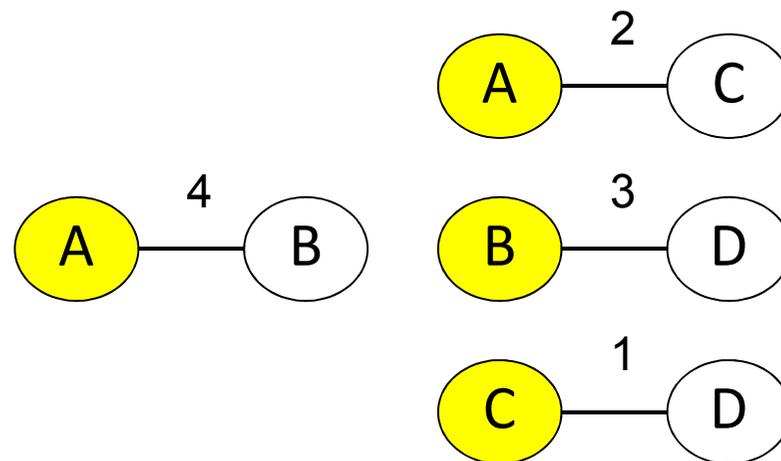
• 興味の対象の薬剤の同種同効薬剤に対する位置づけを明確にしたい

エビデンスを作り出す方法



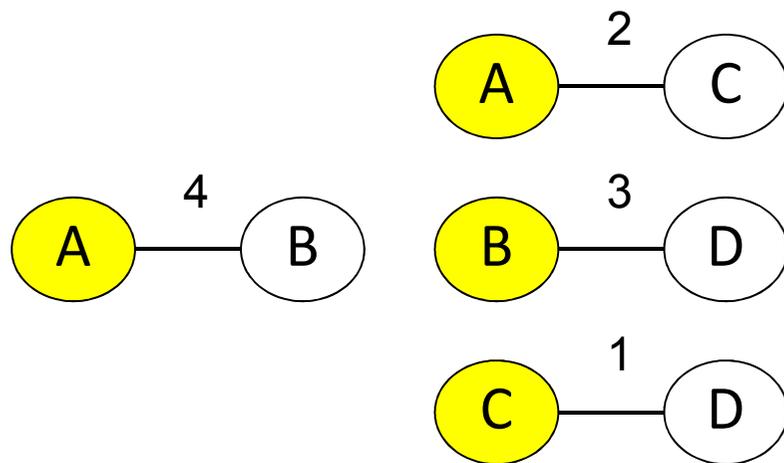
従来型のメタアナリシスの問題点

- 同種同効薬の比較においては...
 - 全ての臨床試験の対照群が同じ薬剤とは限らない
 - 比較対照の変遷(臨床試験で求められる対照群の変遷)
- 対比較のメタアナリシスは...
 - 「対比較」なので、他の同種同効治療があったとしてもメタアナリシスの対象にしない(できない)
 - 先の例: 治療AとBを対象にしたメタアナリシスであり、治療Cに関する試験があってもメタアナリシスの対象にできない

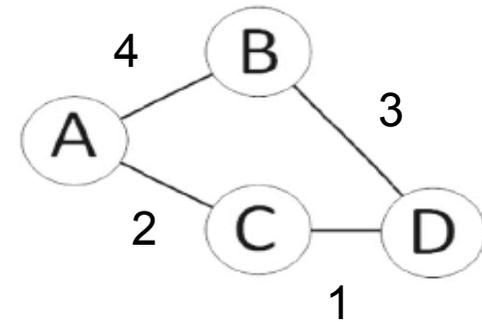
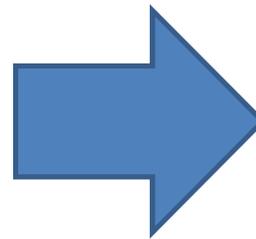


A: 標準治療
黄色: 当該臨床試験での対照治療
数字: 試験数

ネットワークメタアナリシス



A: 標準治療
黄色: 当該試験での対照治療
数字: 試験数



ネットワークメタアナリシス
(Network Meta-analysis: NMA)

NMAの定義

- NMAとは

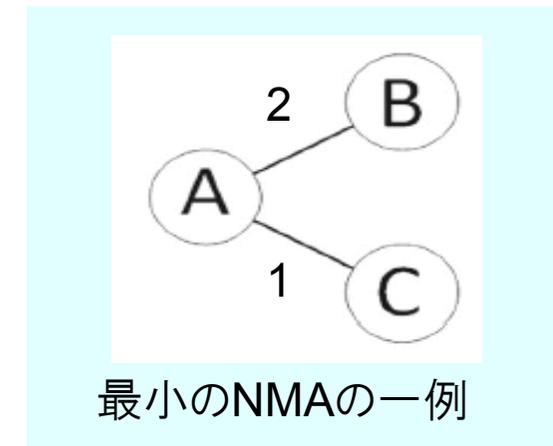
3つ以上の介入をつなぎ, 3つ以上のRCTから構成されるメタアナリシス
[Jansen, et al., 2011]

- エビデンスネットワーク

- 節点: 治療(比較の数を円の大きさと表現することも)
- 線: 治療間の直接的な比較
- 数字: 直接的な比較の数(線の太さで表現することも)

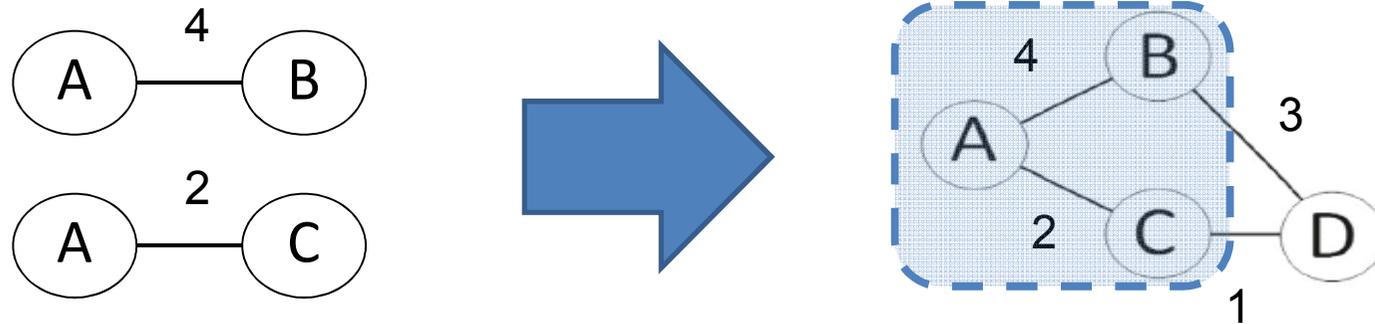
- 基本的にはランダム化比較試験のNMAを想定している

- 観察研究でも可能だが, バイアスの懸念が高まる
 - 観察研究には偏りがつきもの
 - ただでさえ各研究で偏りが懸念されているのに, メタアナリシスで統合したら...
 - 仮に, バイアスの要因が既知だとしても
 - メタアナリシスは各研究で要約した結果を用いるのでバイアスを考慮するのは困難

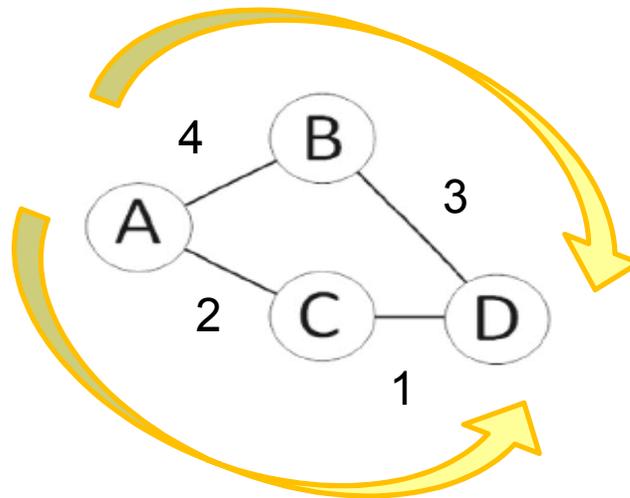


NMAの特徴

- AとB, AとCを個別に比較するのではなく, 「同時に比較」できる



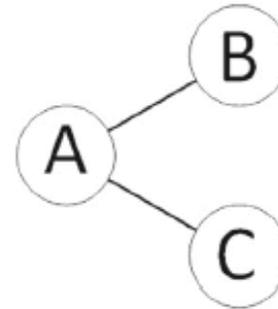
- AとDは直接的に比較できないが, 「間接的に比較」できる
 - Bを介してDと比較
 - Cを介してDと比較



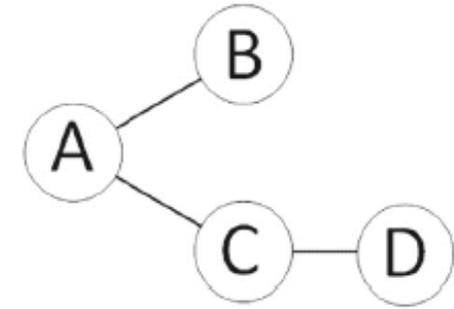
3つの比較とネットワークの様々な形態

- 直接比較

- 比較した試験が有る
- 図例: 線でつながっている比較



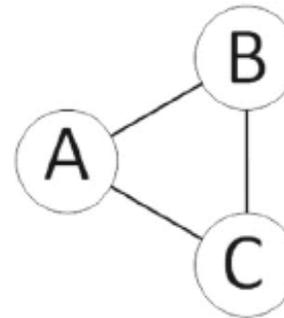
A



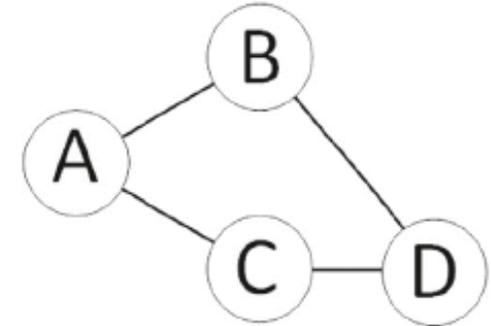
B

- 間接比較

- 比較した試験が無い
- 図例: 右図Aの「BとC」の比較



C



D

- 混合比較

- 直接比較と間接比較ともに含む
- 図例: 右図Cの「BとC」の比較

Fig. 2 – Connected networks of randomized controlled trials to allow for network meta-analysis and indirect comparisons. All interventions that can be indirectly compared are part of one network.

[Jansen, et al., 2014]

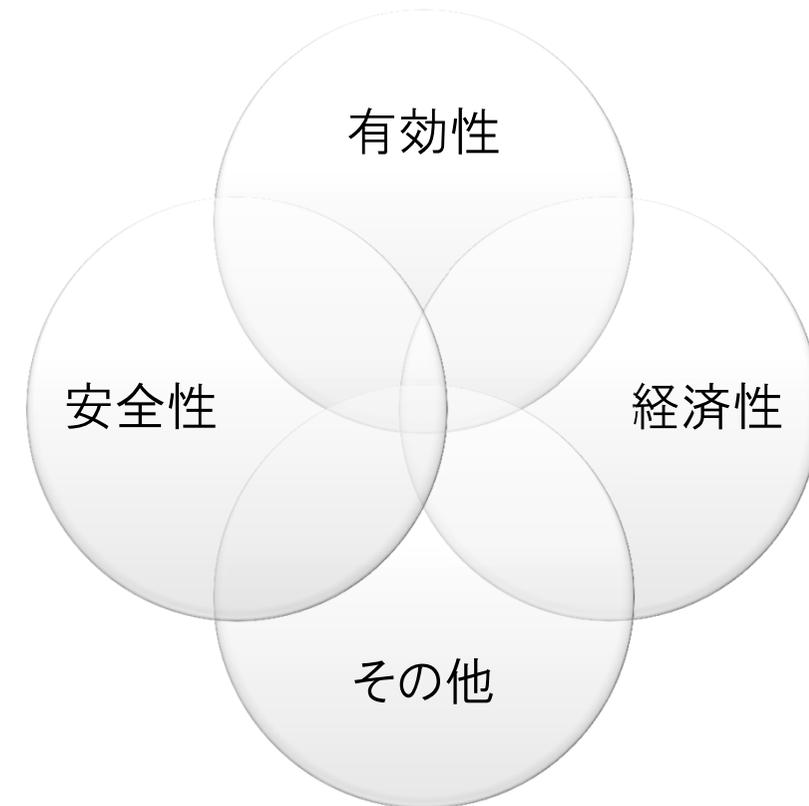
Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis

Leucht S, Cipriani A, Spineli L, et al.

Lancet, 2013; 382(9896): 951-62

NMAの動機

- 統合失調症治療のための抗精神病薬
 - どの薬剤が本疾病の治療に対して好ましいかは意見が分かれる
 - 単に有効性・安全性の面だけではなく、コストも薬剤決定の要素
 - 第二世代の薬剤はコストの面で負担大
- 大目的「新しい薬剤がcost effectiveなのか？」
 - 費用対効果
- 小目的「どの薬剤が有効か？安全か？」
 - 薬剤の序列を決めたい
 - 薬剤が多種なのでNMAを用いる



論文の付録

- プロトコルから結果まで記載された付録(139ページ)

THE LANCET

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Leucht S, Cipriani A, Spineli L, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet* 2013; published online June 27. [http://dx.doi.org/10.1016/S0140-6736\(13\)60733-3](http://dx.doi.org/10.1016/S0140-6736(13)60733-3).

This appendix has been corrected. The corrected version first appeared at thelancet.com on September 13, 2013

Appendix 1 – Protocol (page 2-24)

Appendix 2 – Dose review (page 25-40)

Appendix 3 – Description of included studies (page 41-65)

Appendix 4 – Statistical method (page 66-69)

Appendix 5 – PRISMA statement (page 70-76)

Appendix 6 – Risk of bias assessment (page 77-84)

Appendix 7 – Network of studies for secondary outcomes (page 85-91)

Appendix 8 – Meta-analytic results of direct comparisons (page 92-96)

Appendix 9 – Cumulative ranking curves (page 97-104)

Appendix 10 – Assessment of inconsistency (page 105-114)

Appendix 11 – Sensitivity analyses and meta-regressions (page 115-132)

Appendix 12 – Numbers-needed-to-treat/harm (page 133-139)

NMAの概要

- 対象疾患・治療: 統合失調症の急性治療(治療開始後6週間)
- 対象薬剤: 15種類の抗精神病薬とプラセボ
- 対象試験: 無作為化割付, 盲検
- アウトカム: 有効性, 全ての理由による投与中止, 主な副作用(5種)

一般名	略語	一般名	略語
プラセボ	PBO	クロルプロマジン	CPZ
アセナピン	ASE	ジプラシドン	ZIP
アミスルピリド	AMI	セルチンドール	SER
アリピプラゾール	ARI	ゾテピン	ZOT
イロペリドン	ILO	パリペリドン	PAL
オランザピン	OLA	ハロペリドール	HAL
クエチアピン	QUE	リスペリドン	RIS
クロザピン	CLO	ルラシドン	LUR

黄色は第一世代
水色は第二世代でも比較的新しい

アウトカム

主要

- | | |
|--------------|--|
| ✓ 症状の全体的な変化量 | Positive and Negative Syndrome Scaleのトータルスコアに関するベースラインからの変化量(もしくはBrief Psychiatric Rating Scale) |
|--------------|--|

副次

- | | |
|--|--------------|
| ✓ 全理由による投与中止の有無 | ✓ プロラクチンの増加量 |
| ✓ 体重増加量 | ✓ QTc延長の有無 |
| ✓ 錐体外路系の副作用の有無(代替アウトカムとして抗パーキンソン病薬の使用) | ✓ 鎮静の有無 |

- アウトカムのデータは2人のレビューワーによって独立に抽出
- 可能な限りITT集団のデータを収集
- 欠損などへの対応
 - データの欠損や疑義: 文献の著者に確認
 - 標準偏差の欠損: p値から推定もしくは他試験の標準偏差の平均値から推定

統計解析

• ベイズ流階層モデル(次ページ以降)

- 異質性のパラメータは全ての試験に共有と仮定
- 一貫性: 対比較のメタアナリシスにはないNMAで議論になる性質(次ページ)
 - 一貫性を仮定したモデル(consistencyモデル)
 - 一貫性を仮定しないモデル(inconsistencyモデル)

• 評価指標

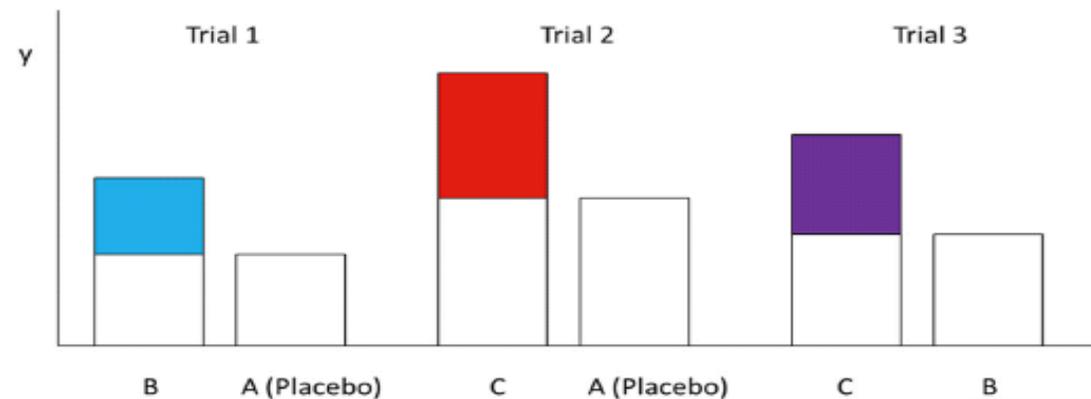
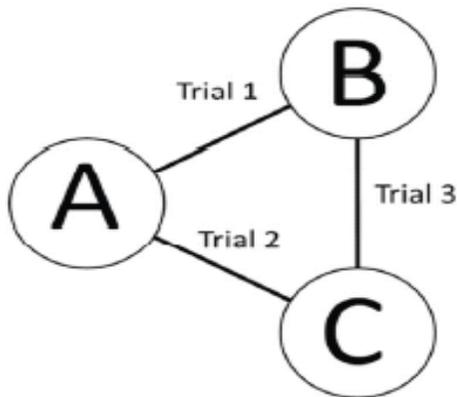
評価指標	
✓ 相対的な指標(いずれも95%信用区間も算出)	
・連続値をとるアウトカム	標準化平均差(Hedges' g)
・二値をとるアウトカム	オッズ比
✓ 順位	SUCRA [Salanti, et al.] (surface under the cumulative ranking)
✓ 絶対的な指標	NNTとNNH

一貫性とは

一貫性の仮定(対義語: 不一致性)

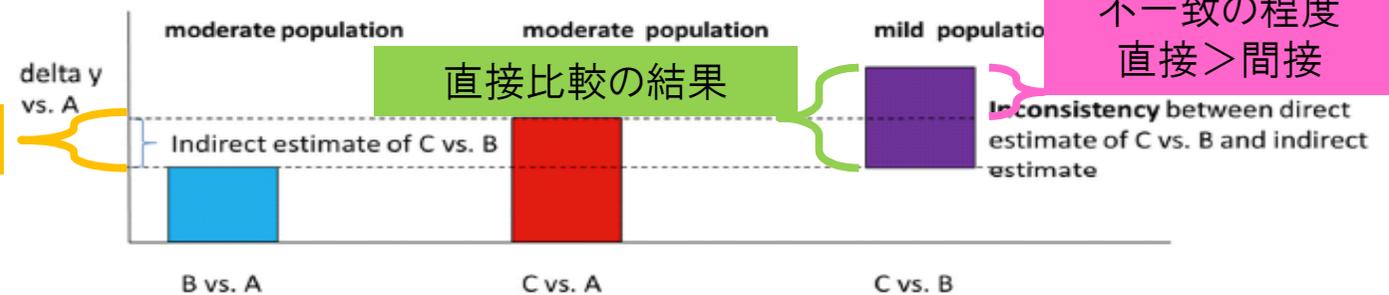
- 直接比較による結果と間接比較のそれが統合するのに十分に一致
- 間接比較にバイアスがなく, 直接・間接間で矛盾がない [Jansen, et al., 2011]

例: BC間の比較に関する不一致性



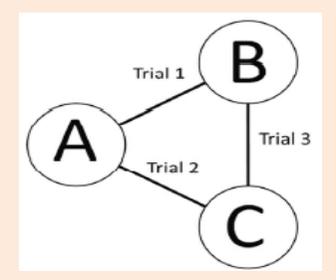
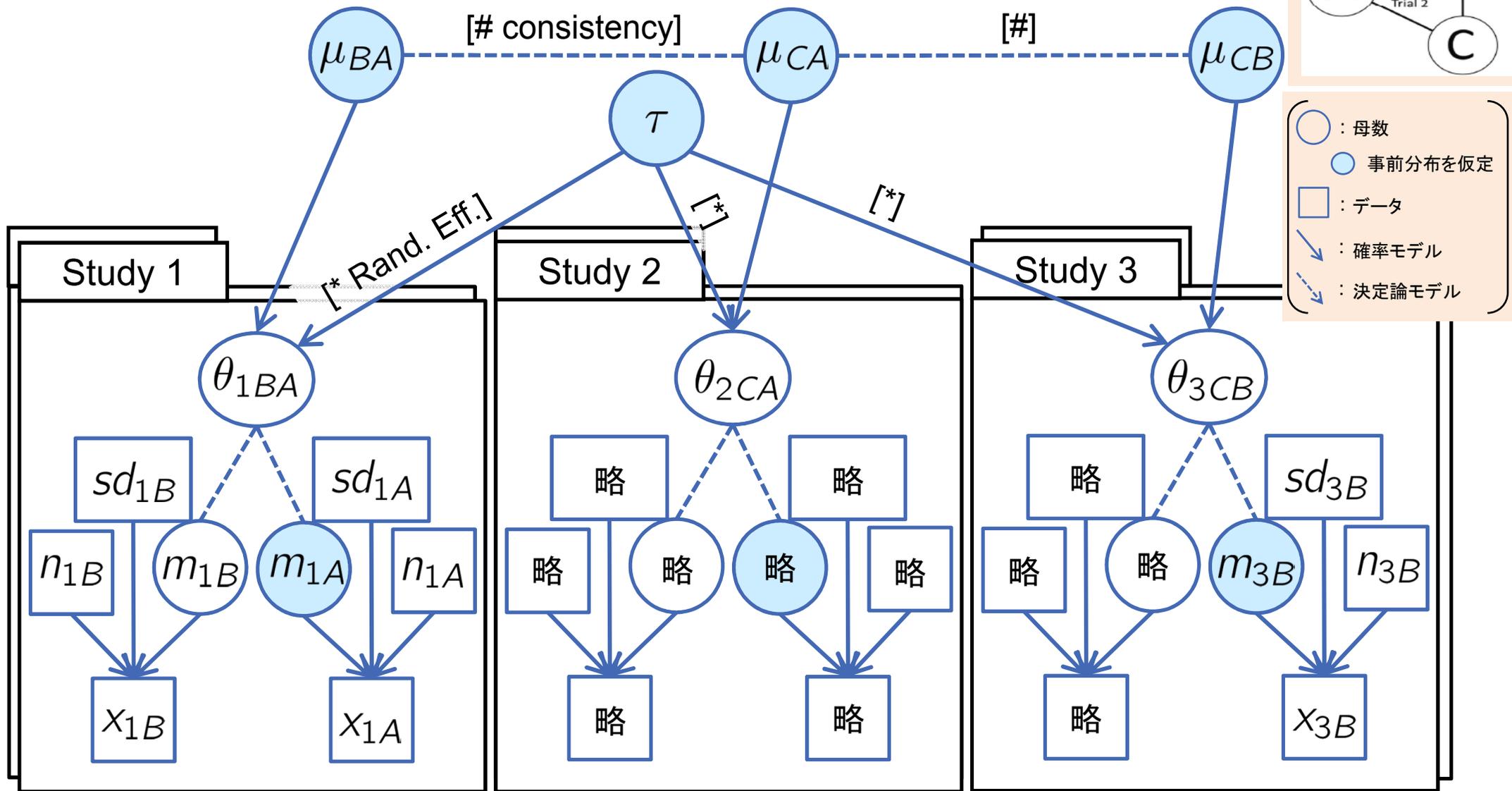
間接比較の結果

Jansen, et al., 2014



不一致の程度
直接 > 間接

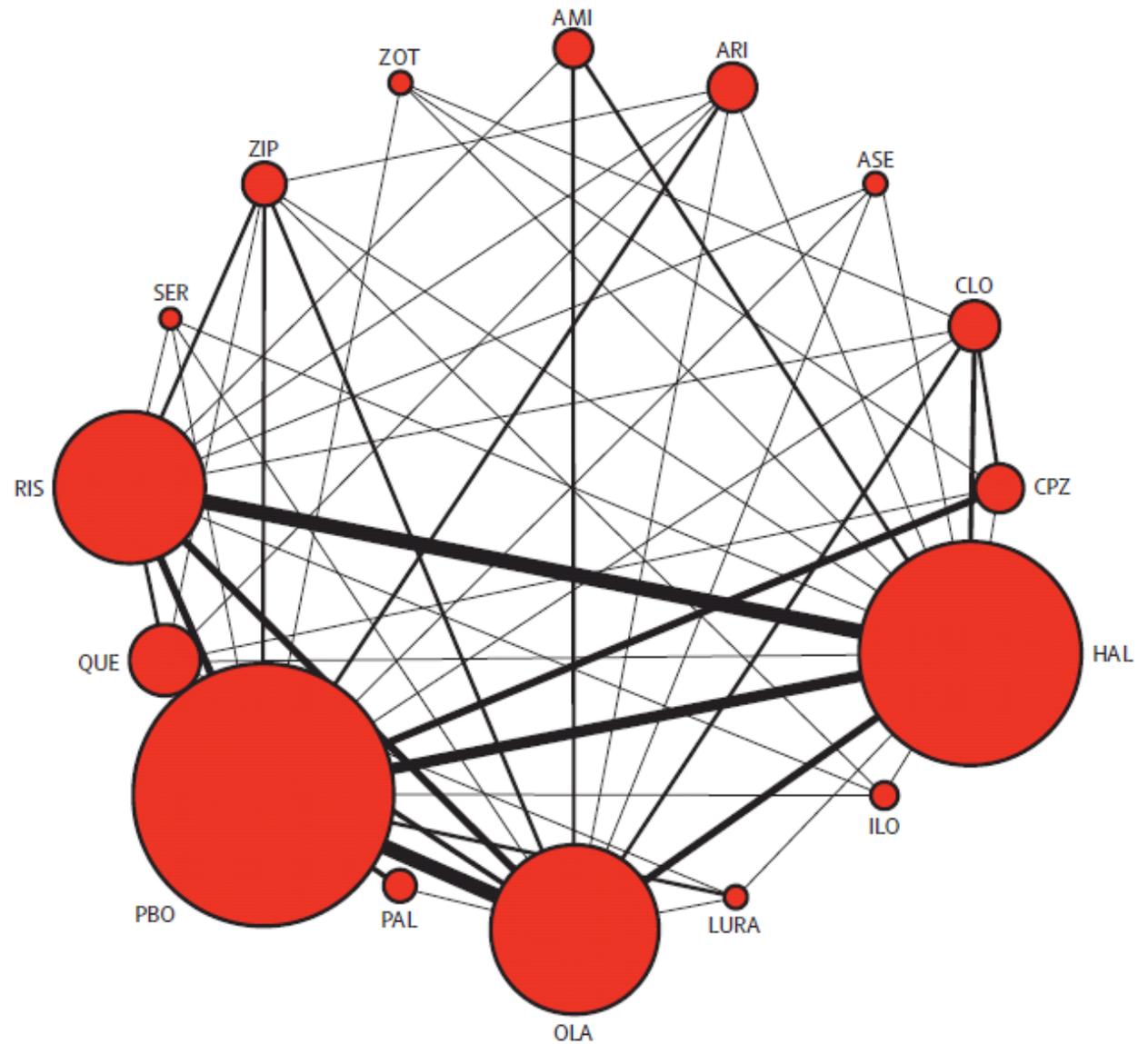
ConsistencyモデルのDAG



ネットワーク図

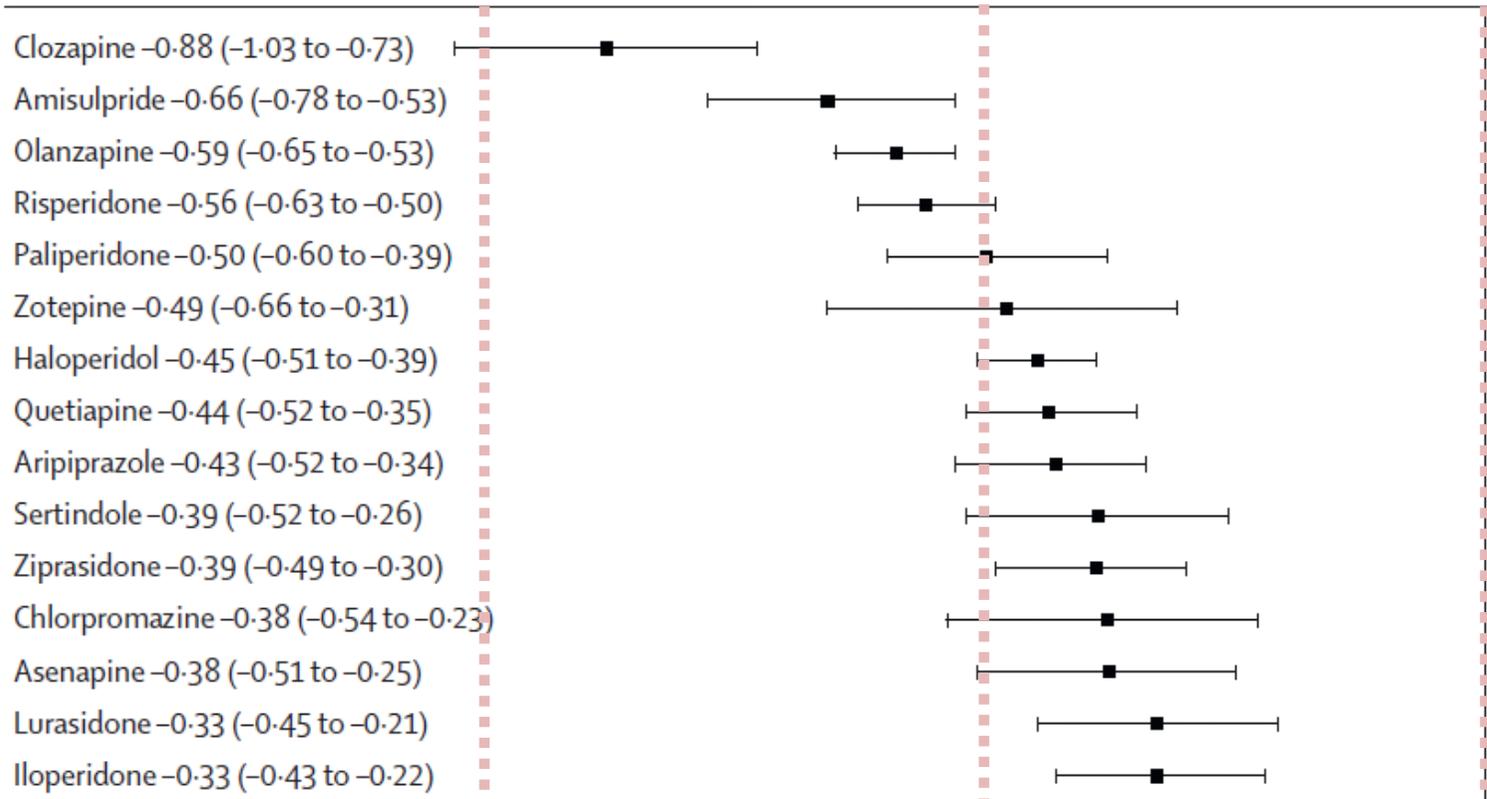
統合失調症治療薬のNMA

- 節点の大きさ: 群数
- 線: 直接比較
- 線の太さ: 試験数に比例



有効性(症状の全体的な変化)の統合結果

<プラセボとの対比較>



- 症状の全体的な変化・・・ベースラインからの変化量は連続値として測定する。
- 標準化平均差・・・各試験の各治療群ごとに平均値が算出されるが、ばらつきを考慮するために、ばらつきで標準化した平均値を求め、これで薬剤間比較をしている。
- 信用区間・・・ベイズ流推定で得られる統計量。「??%の確率で真値がその範囲にある。」

-1 ← 実薬を支持する方向 -0.5 0 プラセボと同等

横軸： 標準化平均差と95%信用区間

<全ての対比較> (左下三角行列が有効性)
 15薬剤 + 1プラセボ → 120通りの比較
 太字は信用区間が0を含まないことを意味している

CLO	1-10	1-00	0-87	0-97	0-70	0-57	0-76	0-76	0-68	0-65	0-71	0-68	0-61	0-67	0-46
-0.22 (-0.41 to -0.04)	AMI	0.93 (0.69 to 1.22)	0.81 (0.60 to 1.08)	0.90 (0.62 to 1.24)	0.66 (0.37 to 1.11)	0.66 (0.58 to 0.76)	0.87 (0.73 to 1.04)	0.88 (0.72 to 1.06)	0.69 (0.53 to 0.88)	0.75 (0.61 to 0.91)	0.83 (0.61 to 1.08)	0.78 (0.60 to 1.01)	0.70 (0.53 to 0.89)	0.78 (0.62 to 0.96)	0.53 (0.46 to 0.60)
-0.29 (-0.44 to -0.14)	-0.07 (-0.19 to 0.05)	OLA	0.87 (0.76 to 1.01)	0.97 (0.78 to 1.20)	0.71 (0.43 to 1.11)	0.66 (0.51 to 1.32)	1.13 (0.66 to 1.78)	1.14 (0.67 to 1.81)	0.90 (0.51 to 1.46)	0.97 (0.56 to 1.55)	1.07 (0.61 to 1.71)	1.02 (0.58 to 1.65)	0.91 (0.51 to 1.47)	1.01 (0.58 to 1.61)	0.69 (0.41 to 1.07)
-0.32 (-0.47 to -0.16)	-0.09 (-0.21 to 0.03)	-0.03 (-0.10 to 0.04)	RIS	1.12 (0.88 to 1.40)	0.82 (0.49 to 1.29)	0.60 (0.48 to 0.75)	0.79 (0.61 to 1.01)	0.79 (0.61 to 1.02)	0.63 (0.46 to 0.85)	0.68 (0.52 to 0.88)	0.75 (0.53 to 1.02)	0.71 (0.52 to 0.95)	0.63 (0.47 to 0.85)	0.70 (0.53 to 0.93)	0.48 (0.39 to 0.58)
-0.38 (-0.57 to -0.20)	-0.16 (-0.32 to -0.00)	-0.09 (-0.21 to 0.02)	-0.07 (-0.19 to 0.06)	PAL	0.74 (0.43 to 1.20)	0.86 (0.51 to 1.32)	1.32 (1.11 to 1.57)	1.33 (1.11 to 1.57)	1.05 (0.82 to 1.31)	1.13 (0.93 to 1.35)	1.25 (0.93 to 1.63)	1.19 (0.92 to 1.50)	1.06 (0.82 to 1.34)	1.17 (0.95 to 1.43)	0.80 (0.71 to 0.90)
-0.39 (-0.60 to -0.19)	-0.17 (-0.38 to 0.04)	-0.10 (-0.29 to 0.08)	-0.08 (-0.26 to 0.11)	0.01 (-0.22 to 0.20)	ZOT	-0.01 (-0.10 to 0.08)	1.01 (0.80 to 1.25)	0.80 (0.60 to 1.04)	0.86 (0.68 to 1.07)	0.95 (0.69 to 1.26)	0.90 (0.68 to 1.19)	0.81 (0.61 to 1.03)	0.89 (0.70 to 1.13)	0.61 (0.52 to 0.71)	
-0.43 (-0.58 to -0.28)	-0.21 (-0.32 to -0.09)	-0.14 (-0.21 to -0.08)	-0.11 (-0.18 to -0.05)	-0.05 (-0.16 to 0.08)	-0.04 (-0.21 to 0.14)	-0.02 (-0.12 to 0.08)	-0.01 (-0.12 to 0.11)	ARI	0.80 (0.59 to 1.04)	0.86 (0.68 to 1.07)	0.95 (0.69 to 1.27)	0.90 (0.68 to 1.18)	0.80 (0.6 to 1.05)	0.89 (0.69 to 1.14)	0.61 (0.51 to 0.72)
-0.44 (-0.61 to -0.28)	-0.22 (-0.36 to -0.08)	-0.15 (-0.25 to -0.06)	-0.13 (-0.22 to -0.03)	-0.06 (-0.19 to 0.08)	-0.05 (-0.24 to 0.14)	-0.01 (-0.10 to 0.08)	1.01 (0.80 to 1.25)	0.80 (0.60 to 1.04)	0.86 (0.68 to 1.07)	0.95 (0.69 to 1.26)	0.90 (0.68 to 1.19)	0.81 (0.61 to 1.03)	0.89 (0.70 to 1.13)	0.61 (0.52 to 0.71)	
-0.45 (-0.62 to -0.28)	-0.23 (-0.37 to -0.08)	-0.16 (-0.25 to -0.07)	-0.13 (-0.23 to -0.03)	-0.07 (-0.20 to 0.08)	-0.06 (-0.25 to 0.14)	-0.02 (-0.12 to 0.08)	-0.01 (-0.12 to 0.11)	SER	0.80 (0.59 to 1.04)	0.86 (0.68 to 1.07)	0.95 (0.69 to 1.27)	0.90 (0.68 to 1.18)	0.80 (0.6 to 1.05)	0.89 (0.69 to 1.14)	0.61 (0.51 to 0.72)
-0.49 (-0.68 to -0.30)	-0.27 (-0.43 to -0.10)	-0.20 (-0.33 to -0.06)	-0.17 (-0.31 to -0.04)	-0.10 (-0.27 to 0.07)	-0.09 (-0.31 to 0.12)	-0.06 (-0.19 to 0.07)	-0.04 (-0.19 to 0.10)	-0.04 (-0.19 to 0.11)	ZIP	1.09 (0.81 to 1.45)	1.21 (0.84 to 1.69)	1.14 (0.81 to 1.56)	1.02 (0.73 to 1.39)	1.13 (0.83 to 1.52)	0.78 (0.61 to 0.98)
-0.49 (-0.66 to -0.31)	-0.26 (-0.41 to -0.12)	-0.20 (-0.29 to -0.10)	-0.17 (-0.27 to -0.07)	-0.10 (-0.24 to 0.04)	-0.09 (-0.29 to 0.11)	-0.05 (-0.15 to 0.04)	-0.04 (-0.16 to 0.08)	-0.04 (-0.16 to 0.09)	-0.15 (-0.15 to 0.16)	CPZ	1.11 (0.80 to 1.50)	1.06 (0.78 to 1.41)	0.94 (0.70 to 1.24)	1.05 (0.81 to 1.33)	0.72 (0.59 to 0.86)
-0.50 (-0.67 to -0.33)	-0.27 (-0.47 to -0.08)	-0.21 (-0.37 to -0.05)	-0.18 (-0.34 to -0.02)	-0.11 (-0.30 to 0.08)	-0.10 (-0.32 to 0.11)	-0.07 (-0.22 to 0.09)	-0.05 (-0.22 to 0.11)	-0.05 (-0.22 to 0.13)	-0.01 (-0.21 to 0.19)	-0.01 (-0.19 to 0.16)	ASE	0.96 (0.66 to 1.34)	0.86 (0.61 to 1.19)	0.96 (0.68 to 1.32)	0.65 (0.50 to 0.84)
-0.50 (-0.69 to -0.30)	-0.27 (-0.45 to -0.10)	-0.21 (-0.34 to -0.08)	-0.18 (-0.32 to -0.04)	-0.11 (-0.28 to 0.05)	-0.10 (-0.32 to 0.11)	-0.07 (-0.20 to 0.07)	-0.05 (-0.20 to 0.09)	-0.05 (-0.20 to 0.10)	-0.01 (-0.19 to 0.17)	-0.01 (-0.17 to 0.14)	0.00 (-0.20 to 0.20)	LUR	0.91 (0.64 to 1.22)	1.01 (0.73 to 1.36)	0.69 (0.54 to 0.86)
-0.55 (-0.74 to -0.36)	-0.33 (-0.50 to -0.16)	-0.26 (-0.39 to -0.13)	-0.23 (-0.37 to -0.10)	-0.17 (-0.33 to -0.00)	-0.16 (-0.37 to 0.06)	-0.12 (-0.25 to 0.01)	-0.11 (-0.25 to 0.03)	-0.10 (-0.25 to 0.05)	-0.06 (-0.24 to 0.11)	-0.07 (-0.22 to 0.09)	-0.05 (-0.25 to 0.14)	-0.05 (-0.23 to 0.12)	ILO	1.12 (0.83 to 1.50)	0.77 (0.61 to 0.96)
-0.55 (-0.73 to -0.38)	-0.33 (-0.48 to -0.18)	-0.26 (-0.38 to -0.15)	-0.24 (-0.35 to -0.12)	-0.17 (-0.32 to -0.02)	-0.16 (-0.36 to 0.04)	-0.12 (-0.23 to -0.02)	-0.11 (-0.24 to 0.02)	-0.10 (-0.24 to 0.03)	-0.07 (-0.23 to 0.10)	-0.07 (-0.20 to 0.06)	-0.06 (-0.24 to 0.13)	-0.06 (-0.22 to 0.11)	0.00 (-0.16 to 0.16)	PBO	0.69 (0.56 to 0.84)
-0.88 (-1.03 to -0.73)	-0.66 (-0.78 to -0.53)	-0.59 (-0.65 to -0.53)	-0.56 (-0.63 to -0.50)	-0.50 (-0.60 to -0.39)	-0.49 (-0.66 to -0.31)	-0.45 (-0.51 to -0.39)	-0.44 (-0.52 to -0.35)	-0.43 (-0.52 to -0.34)	-0.39 (-0.52 to -0.26)	-0.39 (-0.49 to -0.30)	-0.38 (-0.54 to -0.23)	-0.38 (-0.51 to -0.25)	-0.33 (-0.45 to -0.21)	-0.33 (-0.43 to -0.22)	

■ Treatment □ Efficacy (SMD with 95% CrI) □ All cause discontinuation (OR with 95% CrI)

Figure 2: Efficacy and all-cause discontinuation of antipsychotic drugs

NMAの結果に対する視点

- NMAは対比較のMAを単に拡張しただけと考えるのは危険

- 興味の対象を広くし、解析対象の試験が増えたがゆえ・・・

- 試験間の異質性の可能性が大きくなる

類似性のチェック

- 間接比較ができるがゆえ・・・

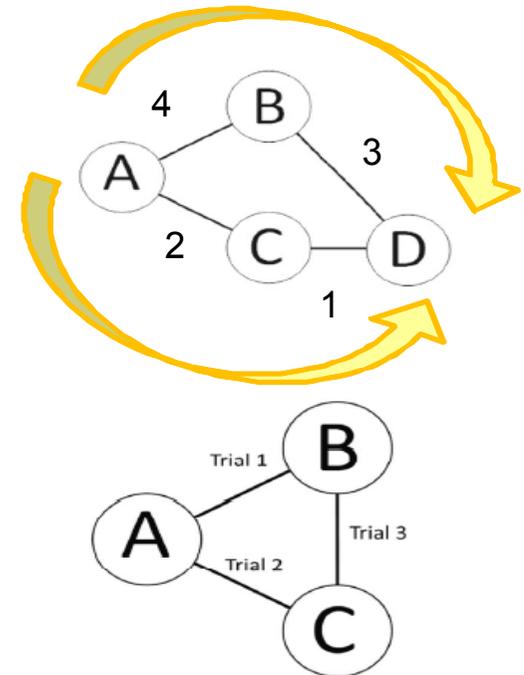
- 間接比較と直接比較の結果の矛盾がおこるかも
- 間接比較は対比較のメタアナリシスにはなかった概念

一致性のチェック

- 自分たちにとって都合のいい結果はそのまま受け止めたくなるが・・・

- NMAの結果を読むときには吟味が必要。それも批判的に

批判的吟味



類似性とその評価

- 類似性(対義語: 異質性)

- NMAを適用する治療群が統合するのに十分に均質であること, つまり各試験のデザイン, 対象集団, およびアウトカムが十分に類似していること [Mills, et al.]
- 満たされないと, NMAの結果にバイアスが生じているかも。解釈要注意

- 類似性の成立および逸脱の影響の主な確認方法

SRの計画段階でのチェック → 類似性の逸脱への未然の対策

- 試験の選択・除外基準の設定, 試験の質の評価
- 定められた方法による文献の収集, 統合可能なアウトカムの選択

SRの実施段階でのチェック → 集まった論文から見る類似性のチェック

- 各試験のRisk of biasを評価し要約する(次ページ)。Riskが高い試験は類似性に影響することが懸念される(感度分析の候補)。

NMAの実施段階でのチェック → 解析結果から見る類似性のチェック

- 各種感度分析(メタ回帰, 部分集団解析)(2, 3ページ後)

Risk of Biasの評価の例

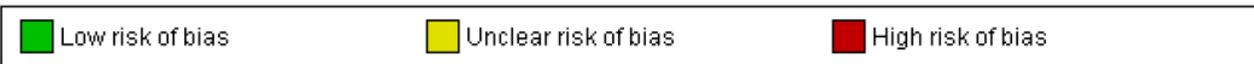
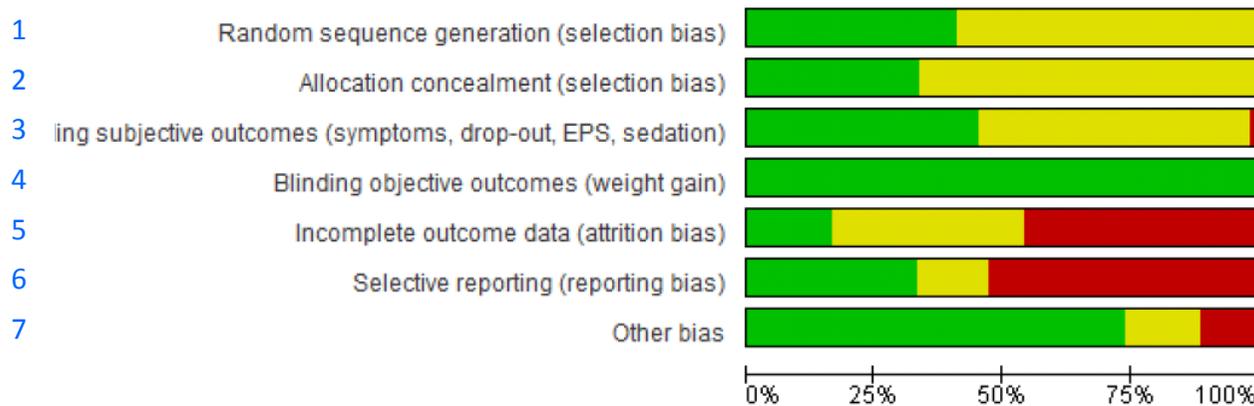
以下7つの観点から各試験を評価

1. 無作為化情報の生成(選択バイアス)
2. 割付の隠匿(選択バイアス)
3. 主観的なアウトカムの隠匿
4. 客観的なアウトカムの隠匿
5. 不完全なアウトカムデータ (attritionバイアス)
6. 選択的な報告(報告バイアス)
7. その他のバイアス

Risk of biasの区分

- Low **+** / Unclear **?** / High **-**

2. 各試験の評価の要約



1. 各試験の評価

	1	2	3	4	5	6	7
Random sequence generation (selection bias)							
Allocation concealment (selection bias)							
Blinding subjective outcomes (symptoms, drop-out, EPS, sedation)							
Blinding objective outcomes (weight gain)							
Incomplete outcome data (attrition bias)							
Selective reporting (reporting bias)							
Other bias							
Addington 2004	+	+	?	+	-	+	+
Adollahian 2008	?	?	?	+	?	-	+
Ahmed 2007	?	?	?	+	-	?	?
Allan 1998	?	?	?	+	?	-	+
Arvanitis 1997	?	?	?	+	-	+	+

各種感度分析

感度分析

- 目的： 異質性と不一致性の潜在的な理由を探索する
- 方法
 - 単盲検試験を除く(n=7)
 - ハロペリドールの試験を除く(n=54)
[第一世代。さまざまな用量の試験が混在している]
 - ITT集団で解析されていない試験を除く(n=18)
 - プラセボを含む試験を除く: プラセボ効果の観点から(n=43)
 - 失敗した試験を除く(n=6)など

メタ回帰

- 目的： 治療と共変量との交互作用を含めた変量(試験レベルの共変量)を導入して、一貫性によるバイアスの影響を減少させる [Cooper, et al.]
- 方法： 共変量を導入
 - 試験期間
 - 脱落率
 - スポンサーシップ
 - 用量
 - 論文の公表年 など

感度分析(一部)

Drug name	全試験の結果		プラセボ 対照試験 を除く(43)	SUCRA rank	失敗した 試験を 除く(6)	SUCRA rank
Clozapine	-0.88 [-1.03, -0.73]	1	-0.51 [-0.69, -0.34]	1	-0.89 [-1.03, -0.75]	1
Amisulpride	-0.66 [-0.78, -0.53]	2	-0.32 [-0.45, -0.19]	2	-0.68 [-0.80, -0.57]	2
Olanzapine	-0.59 [-0.65, -0.53]	3	-0.26 [-0.36, -0.16]	3	-0.63 [-0.68, -0.57]	3
Risperidone	-0.56 [-0.63, -0.50]	4	-0.22 [-0.32, -0.13]	4	-0.58 [-0.64, -0.52]	4
Paliperidone	-0.50 [-0.60, -0.39]	5	-0.20 [-0.36, -0.04]	5	-0.54 [-0.64, -0.45]	5
Zotepine	-0.49 [-0.66, -0.31]	6	-0.18 [-0.39, 0.02]	6	-0.51 [-0.67, -0.34]	6
Haloperidol	-0.45 [-0.51, -0.39]	7	-0.10 [-0.19, -0.01]	9	-0.47 [-0.53, -0.41]	8
Quetiapine	-0.44 [-0.52, -0.35]	8	-0.12 [-0.24, -0.01]	8	-0.47 [-0.55, -0.40]	7
Aripiprazole	-0.43 [-0.52, -0.34]	9	-0.10 [-0.21, 0.02]	10	-0.45 [-0.53, -0.37]	10
Sertindole	-0.39 [-0.52, -0.26]	10	-0.06 [-0.20, 0.08]	11	-0.41 [-0.52, -0.29]	12
Ziprasidone	-0.39 [-0.49, -0.30]	11	-0.06 [-0.16, 0.05]	12	-0.42 [-0.50, -0.34]	11
Chlorpromazine	-0.38 [-0.54, -0.23]	12	0.01 [-0.21, 0.21]	13	-0.40 [-0.54, -0.26]	13
Asenapine	-0.38 [-0.51, -0.25]	13	-0.16 [-0.30, -0.02]	7	-0.47 [-0.58, -0.35]	9
Lurasidone	-0.33 [-0.45, -0.21]	14	0.02 [-0.15, 0.20]	15	-0.39 [-0.51, -0.28]	14
Iloperidone	-0.33 [-0.43, -0.22]	15	<i>Iloperidone = Reference!</i>	14	-0.34 [-0.43, -0.26]	15
Placebo	<i>Reference</i>	16	-	-	<i>Reference</i>	16
Heterogeneity	0.10 [0.06, 0.14]		0.04 [0.01, 0.10]			
Number of studies	167		124			
Heterogeneity	-		↓60%			

概ねロバスト

- 推定値・順位の変動はあるものの
- 例外
 - プラセボ対照試験を除く(43試験)
 - アセナピン
 - 全体の解析第13位, 除くと第7位
 - 失敗に終わった試験を除く(6試験)
 - アセナピン
 - 全体の解析第13位, 除くと第9位

アセナピンの直接比較のペアワイズメタアナリシス

ASE vs. HAL	1	326	-0.02 (-0.25, 0.21)
OLA	3	1647	0.10 (-0.01, 0.21)
PBO	4	942	-0.18 (-0.43, 0.07)
RIS	1	114	-0.15 (-0.51, 0.22)

他の全ての薬剤(プラセボと直接比較のある薬剤)は信用区間が0を含んでいない

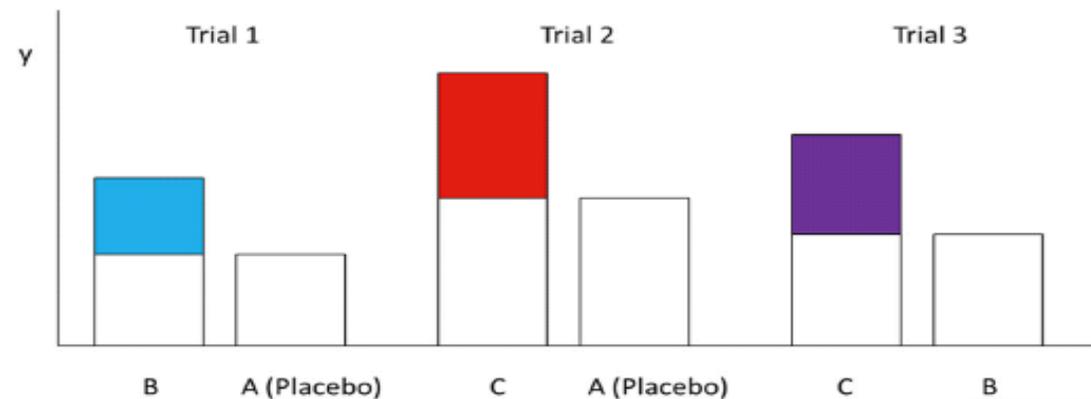
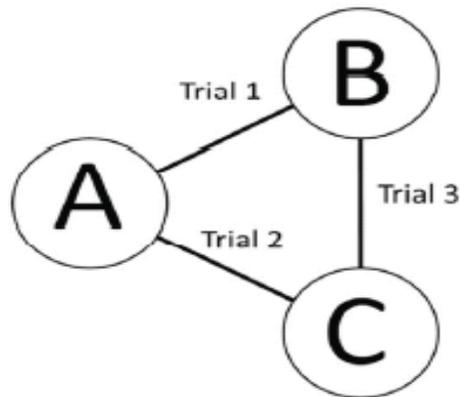
- 異質性
 - プラセボ対照試験を除いたときに60%減

一貫性とは

一貫性の仮定(対義語: 不貫性)

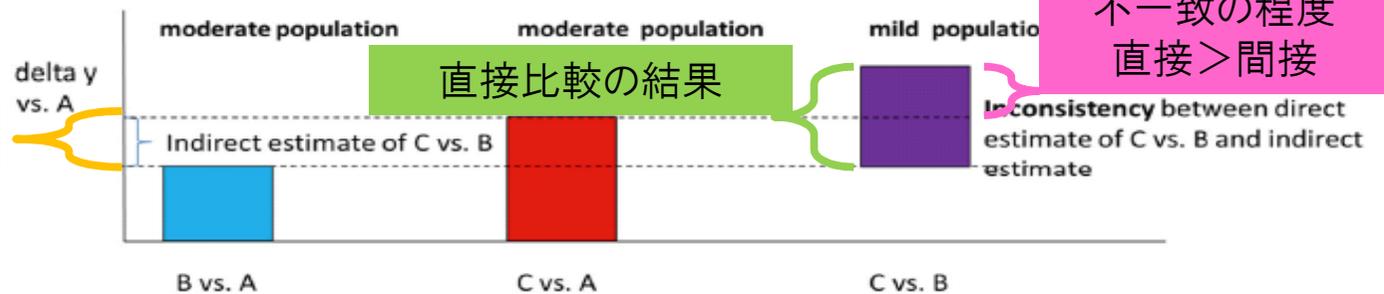
- 直接比較による結果と間接比較のそれが統合するのに十分に一致
- 間接比較にバイアスがなく, 直接・間接間で矛盾がない [Jansen, et al., 2011]

例: BC間の比較に関する不貫性



間接比較の結果

Jansen, et al., 2014



一 致 性 の 評 価

• 一 致 性 の 成 立 お よ び 逸 脱 の 影 響 の 主 な 確 認 方 法 の 例

部分的な一 致 性 の 確 認

- 各 対 比 較 に つ い て 「 間 接 比 較 」 と 「 直 接 比 較 」 を 比 較 す る

全 体 的 な 一 致 性 の 確 認

- 「Consistencyモデル」と「Inconsistencyモデル」とでデータへのあてはまりを比較する

メタアナリシスの段階 → 解析結果から見る一 致 性 の チェック

- 各種感度分析(メタ回帰, 部分集団解析)

一 致 性 の 評 価

- 目的： 間接比較にバイアスがなく、直接・間接間で矛盾がないことを確認する [Jansen et al, 2011]
- 方法：“Inconsistency Factor”
 - Cを介してAとDを比較する場合,

$$IF_{CDA} = |\hat{\mu}_{direct} - \hat{\mu}_{indirect}|$$

$$var(IF_{CDA}) = var(\hat{\mu}_{direct}) + var(\hat{\mu}_{indirect})$$

- ※ 暗に, ループを三角形を成すものに限定している
- 評価方法
 - ループごとの一貫性： 推定値から95%信頼区間を算出し, 区間が0を含むか
 - ネットワーク全体の一貫性： ConsistencyモデルとinconsistencyモデルのDIC
 - 小さい方が当てはまり良好(差が3が基準)

一 致 性 の 評 価 結 果

- 全体の一 致 性 の 評 価 : DICによる評 価

- 小さいほうが好ましい・・・一 致 性 を 仮 定 し た モ デ ル の ほう が 好 ま し い

Model assumption	<i>D</i>	# of data points	<i>DIC</i>
Consistency	403	373	638
Inconsistency	397	373	656

- 部分の一 致 性 の 評 価 : inconsistency factorによる評 価

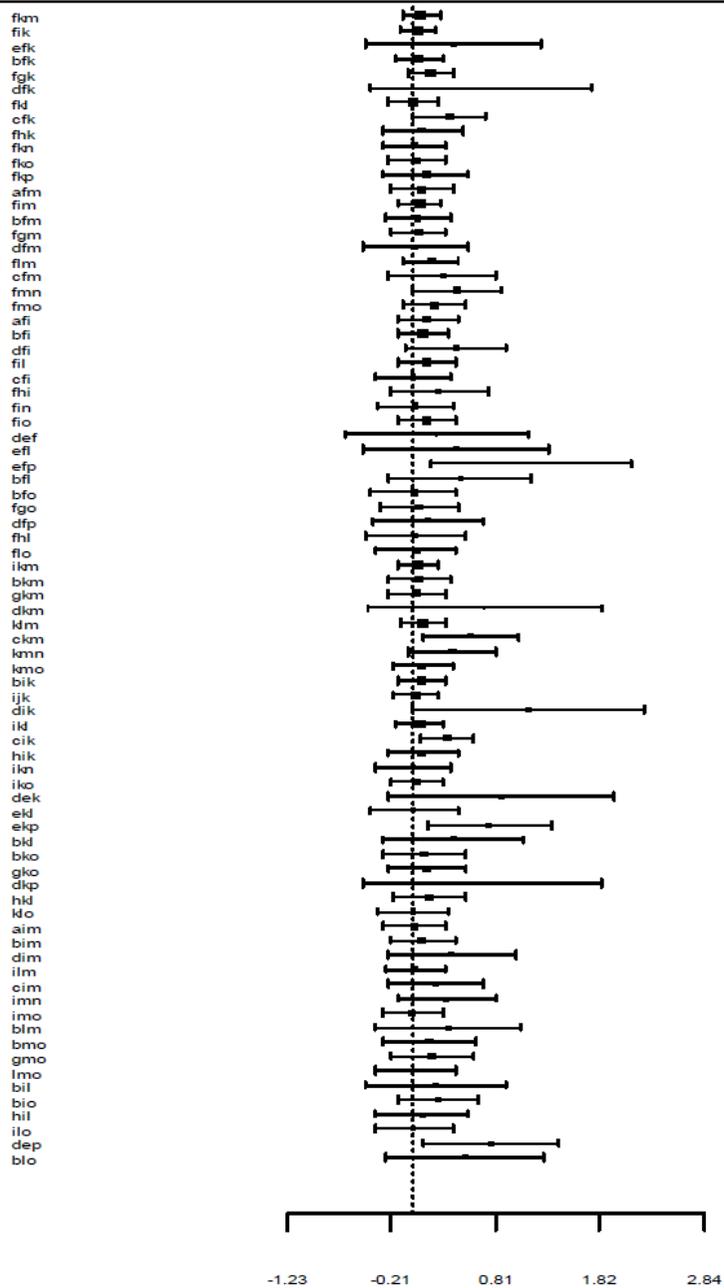
- $IF_{CDA} = |\hat{\mu}_{direct} - \hat{\mu}_{indirect}|$

- 次ページ

- 信 頼 区 間 が 0 を 含 ま な い 比 較 の 個 数 ・ ・ ・ 80 個 の ル ー プ 中 7 個 (約 9 %) By chance?

- 結 論

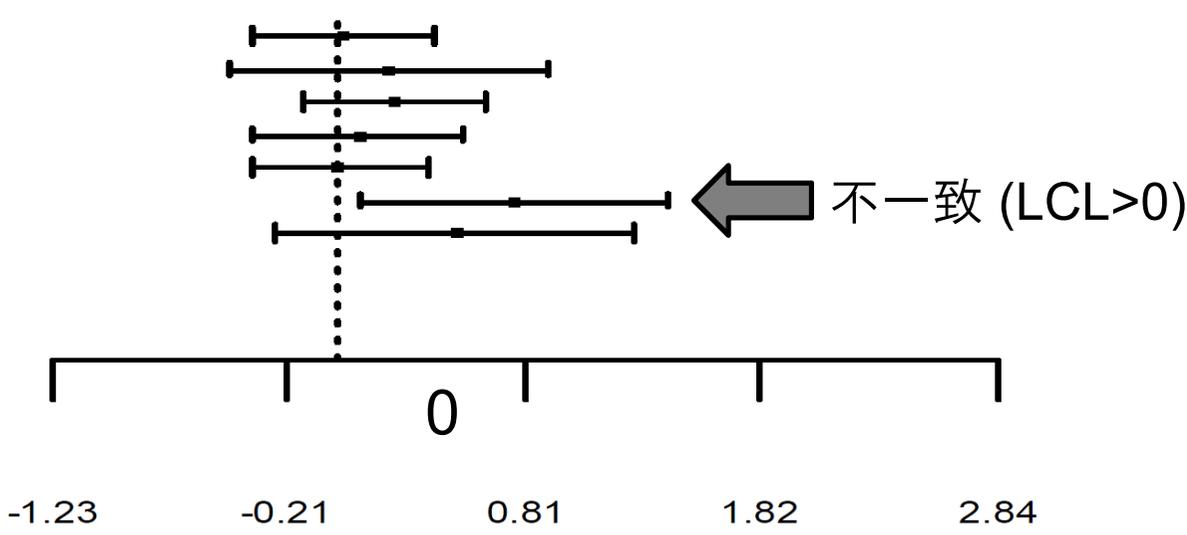
- 不 一 致 性 が 確 認 さ れ た も の の 多 く が そ の 程 度 は 小 さ く , 結 果 を 覆 す ほ ど で も な い



```

0.08 [-0.10, 0.27]
0.05 [-0.13, 0.24]
0.40 [-0.45, 1.26]
0.07 [-0.18, 0.31]
0.17 [-0.06, 0.40]
0.67 [-0.40, 1.73]
0.00 [-0.24, 0.25]
0.36 [-0.01, 0.72]
0.09 [-0.29, 0.48]
0.02 [-0.28, 0.31]
0.04 [-0.25, 0.33]
0.14 [-0.28, 0.55]
0.09 [-0.21, 0.39]
0.08 [-0.13, 0.28]
0.05 [-0.26, 0.37]
0.05 [-0.22, 0.33]
0.04 [-0.48, 0.55]
0.18 [-0.08, 0.44]
0.29 [-0.24, 0.82]
0.43 [0.01, 0.85]
0.20 [-0.10, 0.51]
0.15 [-0.14, 0.44]
0.10 [-0.15, 0.36]
0.42 [-0.07, 0.90]
0.13 [-0.15, 0.41]
0.01 [-0.36, 0.38]
0.26 [-0.22, 0.73]
0.03 [-0.33, 0.39]
0.13 [-0.15, 0.41]
0.23 [-0.65, 1.12]
0.42 [-0.49, 1.32]
1.15 [0.18, 2.12]
0.48 [-0.23, 1.16]
0.01 [-0.40, 0.43]
0.07 [-0.31, 0.45]
0.15 [-0.39, 0.69]
0.03 [-0.47, 0.53]
0.04 [-0.35, 0.42]
0.06 [-0.13, 0.25]
0.06 [-0.25, 0.36]
0.04 [-0.24, 0.32]
0.69 [-0.45, 1.84]
0.11 [-0.11, 0.33]
0.57 [-0.10, 1.23]

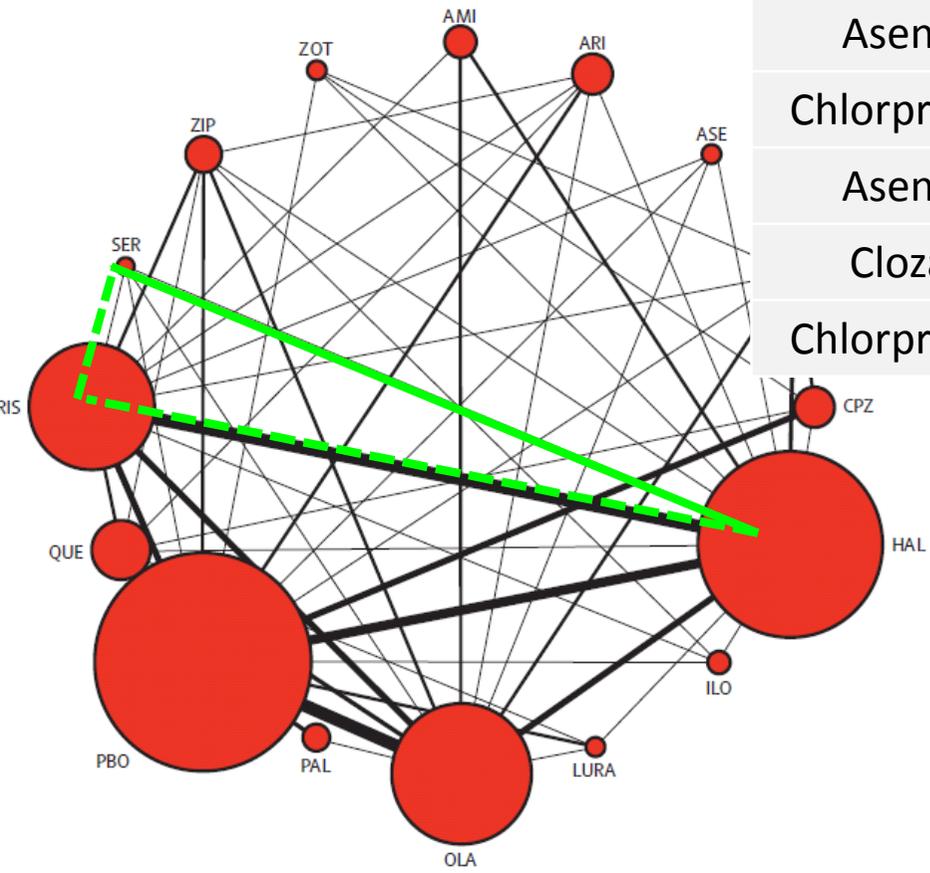
```



対照	アンカー	興味	差	95%CI
Haloperidol	Risperidone	Sertindole	0.43	(0.01, 0.85)
Clozapine	Haloperidol	Zotepine	1.15	(0.18, 2.12)
Asenapine	Placebo	Risperidone	0.57	(0.10, 1.03)
Chlorpromazine	Olanzapine	Placebo	1.13	(0.01, 2.26)
Asenapine	Olanzapine	Placebo	0.34	(0.08, 0.60)
Clozapine	Placebo	Zotepine	0.75	(0.16, 1.34)
Chlorpromazine	Clozapine	Zotepine	0.76	(0.11, 1.42)

-1.23 -0.21 0.81 1.82 2.84
Inconsistency for triangular loops

対照	アンカー	興味	差	95%CI
Haloperidol	Risperidone	Sertindole	0.43	(0.01, 0.85)
Clozapine	Haloperidol	Zotepine	1.15	(0.18, 2.12)
Asenapine	Placebo	Risperidone	0.57	(0.10, 1.03)
Chlorpromazine	Olanzapine	Placebo	1.13	(0.01, 2.26)
Asenapine	Olanzapine	Placebo	0.34	(0.08, 0.60)
Clozapine	Placebo	Zotepine	0.75	(0.16, 1.34)
Chlorpromazine	Clozapine	Zotepine	0.76	(0.11, 1.42)



- 節点の大きさ・・・群数
- 線・・・直接比較されている試験
- 線の太さ・・・試験数

- 試験数が多い比較(線が太い)を含むグループでも不一致が起こっている(4グループ)
 - 例: Haloperidol, Risperidone, Sertindole
 - 直接比較
 - Haloperidol ⇔ Sertindole (3試験)
 - 間接比較
 - Haloperidol ⇔ Risperidone (16試験)
 - Risperidone ⇔ Sertindole (1試験)

批判的吟味

- NMAの結果を読むときには吟味が必要。それも批判的に
 - 自分たちにとって魅力のある結果はそのまま受け止めてしまいたくなる
 - 自分たちにとって不都合な結果は蓋をして追いやってしまいたくなる

 - 何れの場合も、目の前のNMAの結果が妥当か批判的な立場で見るべき
 - 著名な論文誌の論文だから信頼できる？…そうとは限らないかも
 - 自分たちのresearch question(PECO/PICO)と照らして参考に値するか？
 - 例： 対象治療の構成, 対象集団, 試験実施の時期, エンドポイント

 - 批判的吟味の材料(後述)
 - SRのプロトコール
 - NMAの妥当性や信憑性の評価のガイドとしてのチェックリスト
[Jansen, et al., 2011, Jansen, et al., 2014, Mills, et al.]

システマティックレビューのプロトコール

- プロトコールの事前提出と公表の義務化？

- PROSPERO
- 論文の付録としてプロトコールを提示
 - 例： 統合失調症のNMAの付録(プロトコール以外も含む)

THE LANCET

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Leucht S, Cipriani A, Spineli L, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet* 2013; published online June 27. [http://dx.doi.org/10.1016/S0140-6736\(13\)60733-3](http://dx.doi.org/10.1016/S0140-6736(13)60733-3).

This appendix has been corrected. The corrected version first appeared at thelancet.com on September 13, 2013

[Leucht, et al.]

- Appendix 1 – Protocol (page 2-24)
- Appendix 2 – Dose review (page 25-40)
- Appendix 3 – Description of included studies (page 41-65)
- Appendix 4 – Statistical method (page 66-69)
- Appendix 5 – PRISMA statement (page 70-76)
- Appendix 6 – Risk of bias assessment (page 77-84)
- Appendix 7 – Network of studies for secondary outcomes (page 85-91)
- Appendix 8 – Meta-analytic results of direct comparisons (page 92-96)
- Appendix 9 – Cumulative ranking curves (page 97-104)
- Appendix 10 – Assessment of inconsistency (page 105-114)
- Appendix 11 – Sensitivity analyses and meta-regressions (page 115-132)
- Appendix 12 – Numbers-needed-to-treat/harm (page 133-139)

- 「SRのプロトコール」と「論文を読む側の目的」と照らす

- Research question, PECO/PICO

NMAの妥当性や信憑性の評価のガイドとしてのチェックリスト

Jansen, et al., 2014

Relevance

- Is the population relevant?
- Are any relevant interventions missing?
- Are any relevant outcomes missing?
- Is the context (settings and circumstances) applicable?

Credibility

Evidence Base Used for the Indirect Comparison or Network Meta-Analysis

- Attempt to include all relevant RCTs?
- 1 network?
- No poor quality RCTs?
- No differences in effect modifiers between direct comparisons?

Analysis

- Naiive comparisons avoided?
- Consistency assessed?
- With consistency, was direct & indirect evidence included?
- Account for inconsistency/Minimize bias?
- Valid rationale for FE/RE model?
- Rationale for hetelogeneity assumptions in RE model discussed?
- Subgroup or meta-regression analysis?

Reporting Quality and Transparency

- Network & source data presented?
- Direct & indirect results reported?
- Are all contrasts presented with uncertainty?
- Ranking of treatments presented?
- Results by subgroup or levels of effect-modifiers presented?

Interpretation

- Conclusions fair & balanced?

Conflict of Interest

- Conflict of interest? If yes, steps taken to address these?

最後に

- NMAは治療比較における強力なツール
 - ある意味, 究極的な要求を満たしてくれる(どの治療が一番か?)
- NMAは甘い毒林檎? パンドラの箱??
 - NMAは対比較のメタアナリシスよりも不確実な要素が多い
 - 類似性⇔異質性: 対比較のメタアナリシスよりも対象が広い
 - 一貫性⇔不一致性: 対比較のメタアナリシスには無かった仮定
 - NMAの結果を見る立場として, 妥当性を十分に確認: 批判的吟味
 - たとえ結果が自分にとってポジな結果であったとしても
 - 立ち止まって目の前のNMAを疑って
 - 自分たちのresearch questionと照らして参考に値するか?
- 客観的視点とSRやNMAのリテラシー
 - MA, 臨床, 統計などの各専門知識を総動員しなければならない

引用文献

- Carroll K, et al. On the need for increased rigour and care in the conduct and interpretation of network meta-analyses in drug development. *Pharm Stat*, doi: 10.1002/pst.1733 (early view).
- Cooper NJ, et al. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: application to stroke prevention treatments in individuals with nonrheumatic atrial fibrillation. *Stat Med*, 2009 **28**, 1861–81.
- Jansen JP, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. *Value Health*, 2011. **14**, 417-28.
- Jansen JP, et al. Indirect Treatment Comparison/Network Meta-Analysis Study Questionnaire to Assess Relevance and Credibility to Inform Health Care Decision Making: An ISPOR-AMCP-NPC Good Practice Task Force Report. *Value Health*, 2014. **17**, 157-73.
- Leucht S, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet*, 2013. **382**, 951-62.
- Mills EJ, et al. How to use an article reporting a multiple treatment comparison meta-analysis. *Journal of the American Medical Association*, 2012. **29** 1246-53.
- Salanti G, et al. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol*, 2011. **64**: 163–71.

話題提供 野間先生

内容

- 一貫性・不一貫性の評価方法
 - エスシタロプラムとシタロプラムで不一貫性。結果, 検定($p=0.009$)
 - 一つの比較で一貫性が崩れると全体の比較に影響する
 - 直接比較・・・Industry sponsorship bias
 - 間接比較・・・
- FDALレジストレーション
 - Publication bias
- メタアナとその後の大規模試験
 - これらの食い違い
- チェックリスト
 - PRISMA for NMA (2015)
- STATAによる解析ツール
 - By Cochrane group
 - Inconsistency, バイアスの評価も可能
- NMAは“Epidemiology of clinical trials”

グループディスカッション1：話題提供の振り返り - 質問コーナー -

藤井 陽介

グループディスカッション1:追加の説明
- 事例関連 -
[Leucht, et al.]

藤井 陽介

プロトコールの一部

疾患・治療・薬剤の選択と除外

対象とした疾病・状態	除外した疾病・状態(試験間で試験の特性を均一にするため [9, 10])
<ul style="list-style-type: none"> 統合失調症 それに関連した疾病 <ul style="list-style-type: none"> 統合失調感情障害 統合失調症様障害 妄想性障害 	<ul style="list-style-type: none"> 陰性症状が優勢 併発している疾患 治療抵抗性 持続した症状を有する患者
対象とした治療	対象とした薬剤
<ul style="list-style-type: none"> 急性治療(急性は投与開始後6週)。6週データがない場合は4週もしくは12週。 	<ul style="list-style-type: none"> 15種類の単独経口投与の抗精神病薬 増量を認めた試験も含む※

試験の質の評価

組み入れた試験	除外した試験
<ul style="list-style-type: none"> 出版されていないものも含め、少なくとも単盲検であるような試験を含めた 	<ul style="list-style-type: none"> バイアスの恐れがあるようなsequence generationを含むような試験 そして割付が確実に盲検されていない(例えばalternate allocation)試験 非盲検試験 <ul style="list-style-type: none"> 以前の解析において第二世代の薬剤に系統的に良好な結果となったため [3] 中国で実施された試験 無作為化手順が適切でない試験や手順を報告していない試験が数多く存在したため [23] 投与群のクロスオーバーが可能な試験
<ul style="list-style-type: none"> 評価は二人のレビューワーによって独立に実施 	

文献の収集

手順1	<ul style="list-style-type: none"> 過去に行われた7つのシステマティックレビューで同定された文献を照合 [3, 6, 15-19]
手順2	<ul style="list-style-type: none"> 2012年9月1日までに公表された文献を対象として検索 <ul style="list-style-type: none"> Cochrane Schizophrenia Group's specialised register Medline など 文献に掲載の結果をUS FDAのレポートや製薬企業への確認によって補填 検索語も明確に規定(抗精神病薬の一般名やエンドポイント名)
手順3	<ul style="list-style-type: none"> US FDAや製薬企業のウェブサイトに掲載されているレポートを確認しリスト化。

アウトカムの選択

主要	
✓ 症状の全体的な変化量	Positive and Negative Syndrome Scale [25] のトータルスコアに関するベースラインからの変化量(もしくはBrief Psychiatric Rating Scale [26])
副次	
✓ 全理由による投与中止の有無	✓ プロラクチンの増加量
✓ 体重増加量	✓ QTc延長の有無
✓ 錐体外路系の副作用の有無(代替アウトカムとして抗パーキンソン病薬の使用)	✓ 鎮静の有無
<ul style="list-style-type: none"> アウトカムのデータは2人のレビューワーによって独立に抽出 可能な限りITT集団のデータを収集 欠損などへの対応 <ul style="list-style-type: none"> データの欠損や疑義: 文献の著者に確認 標準偏差の欠損: p値から推定もしくは他試験の標準偏差の平均値から推定 [27] 	

疾患・治療・薬剤の選択

対象とした疾患・状態

- 統合失調症
- それに関連した疾病
 - 統合失調感情障害
 - 統合失調症様障害
 - 妄想性障害

除外した疾患・状態(試験間で試験の特性を均一にするため)

- 陰性症状が優勢
- 併発している疾患
- 治療抵抗性
- 持続した症状を有する患者

対象とした治療

- 急性治療(急性は投与開始後6週)。6週データがない場合は4週もしくは12週。

対象とした薬剤

- 15種類の単独経口投与の抗精神病薬
- 増量を認めた試験も含む※

※用量についてコンセンサスをとるために行われた試験から定められた用量を基に、各試験の最大用量と照らし合わせて組み入れた

文献の収集

手順1

- 過去に行われた7つのシステマティックレビューで同定された文献を照合

手順2

- 2012年9月1日までに公表された文献を対象として検索
 - Cochrane Schizophrenia Group's specialised register
 - Medlineなど
- 文献に掲載の結果をUS FDAのレポートや製薬企業への確認によって補填
- 検索語も明確に規定(抗精神病薬の一般名やエンドポイント名)

手順3

- US FDAや製薬企業のウェブサイトに掲載されているレポートを確認しリスト化

試験の質の評価

組み入れた試験

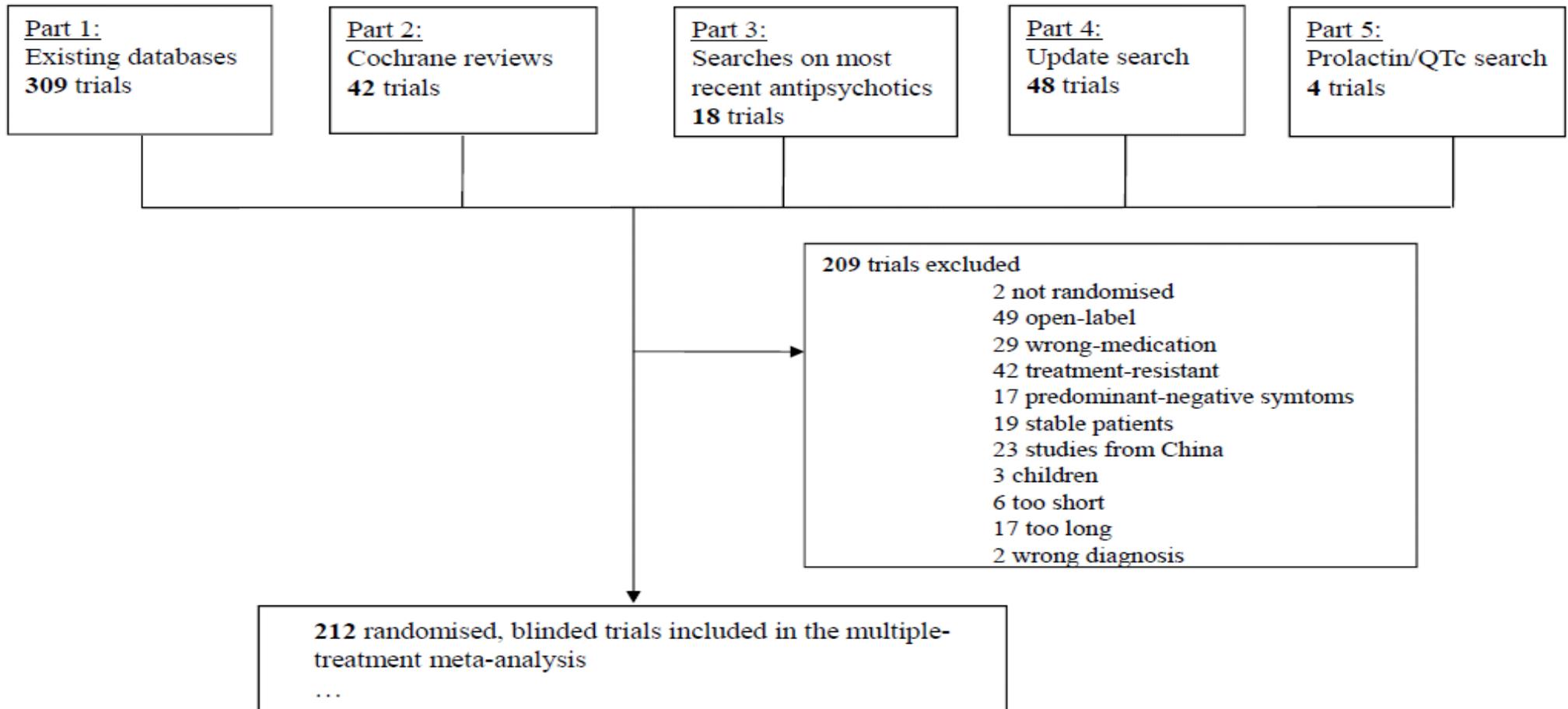
- 出版されていないものも含め、少なくとも単盲検であるような試験を含めた

除外した試験

- バイアスの恐れがあるようなsequence generationを含むような試験
- そして割付が確実に盲検されていない(例えばalternate allocation)試験
- 非盲検試験
 - 以前の解析において第二世代の薬剤に系統的に良好な結果となったため
- 中国で実施された試験
 - 無作為化手順が適切でない試験や手順を報告していない試験が数多く存在したため
- 投与群のクロスオーバーが可能な試験

- 評価は二人のレビューワーによって独立に実施

システマティックレビュー： PRISMAダイアグラム



メタ回帰(一部)

- 試験規模に関するメタ回帰
 - 各試験の標準誤差で調整
 - 右の表は最も大きな規模(最も小さな標準誤差)のプラセボ対照試験の標準誤差としたときの標準化平均差
- 結果: プラセボに対する有効性
 - 小規模な試験 > 大規模な試験
 - 小規模の試験では実薬の有効性を過大評価してしまう傾向を明らかにした (coefficient = 1.31, 95% CrI (0.58-2.03))
 - しかしながら順位への影響は小

Drug name	Standard analysis	Rank	Adjusted for small trial effects	Rank
Clozapine	-0.88 [-1.03, -0.73]	1	-0.69 [-0.87, -0.50]	1
Amisulpride	-0.66 [-0.78, -0.53]	2	-0.51 [-0.65, -0.37]	2
Olanzapine	-0.59 [-0.65, -0.53]	3	-0.45 [-0.55, -0.35]	3
Risperidone	-0.56 [-0.63, -0.50]	4	-0.42 [-0.52, -0.32]	4
Paliperidone	-0.50 [-0.60, -0.39]	5	-0.36 [-0.49, -0.24]	5
Zotepine	-0.49 [-0.66, -0.31]	6	-0.31 [-0.51, -0.11]	6
Haloperidol	-0.45 [-0.51, -0.39]	7	-0.30 [-0.40, -0.19]	7
Quetiapine	-0.44 [-0.52, -0.35]	8	-0.30 [-0.41, -0.18]	8
Aripiprazole	-0.43 [-0.52, -0.34]	9	-0.30 [-0.42, -0.18]	9
Sertindole	-0.39 [-0.52, -0.26]	10	-0.25 [-0.40, -0.09]	10
Ziprasidone	-0.39 [-0.49, -0.30]	11	-0.25 [-0.38, -0.13]	11
Chlorpromazine	-0.38 [-0.54, -0.23]	12	-0.12 [-0.33, 0.09]	15
Asenapine	-0.38 [-0.51, -0.25]	13	-0.25 [-0.40, -0.10]	12
Lurasidone	-0.33 [-0.45, -0.21]	14	-0.21 [-0.34, -0.07]	13
Iloperidone	-0.33 [-0.43, -0.22]	15	-0.20 [-0.32, -0.07]	14
Placebo	<i>Reference</i>	16	<i>Reference</i>	16

SUCRA (surface under the cumulative ranking)

- 順位付けのための指標

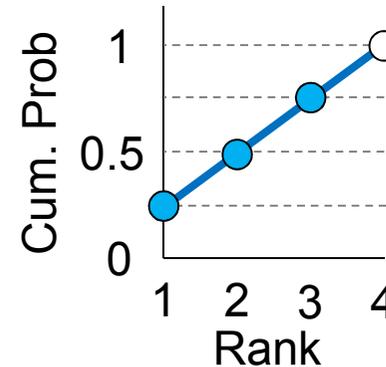
- ランキングプロット
- SUCRA [Salanti et al.]

$$SUCRA_j = \frac{\sum_{b=1}^{a-1} cum_{j,b}}{a-1}$$

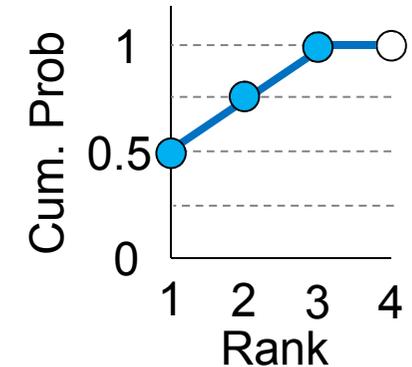
a: 薬剤数
 b: 順位
 j: 治療
 cum_{j,b}: 累積確率

Rank <i>b</i>	Treatment <i>j</i>			
	A	B	C	D
1	0.25 (0.25)	0.50 (0.50)	0.25 (0.25)	0 (0)
2	0.25 (0.50)	0.25 (0.75)	0.50 (0.75)	0 (0)
3	0.25 (0.75)	0.25 (1)	0.25 (1)	0.25 (0.25)
4	0.25 (1)	0 (1)	0 (1)	0.75 (1)
<i>SUCRA</i>	50%	75%	67%	8%

Treatment A



Treatment B



- SUCRAの解釈

- 100% …… どの薬剤と比較しても完全に優れている治療
- 0% …… どの薬剤と比較しても完全に劣っている治療
- X% …… どの薬剤にも完全に優れている治療の有効性のX%を達成することを意味する

順位付け

• 順位付け

- NMAの解析の一手法に、興味のエンドポイントに関して良い(悪い)順に順位を付けるものがある (Bayes流の場合)

• 順位付けのための指標

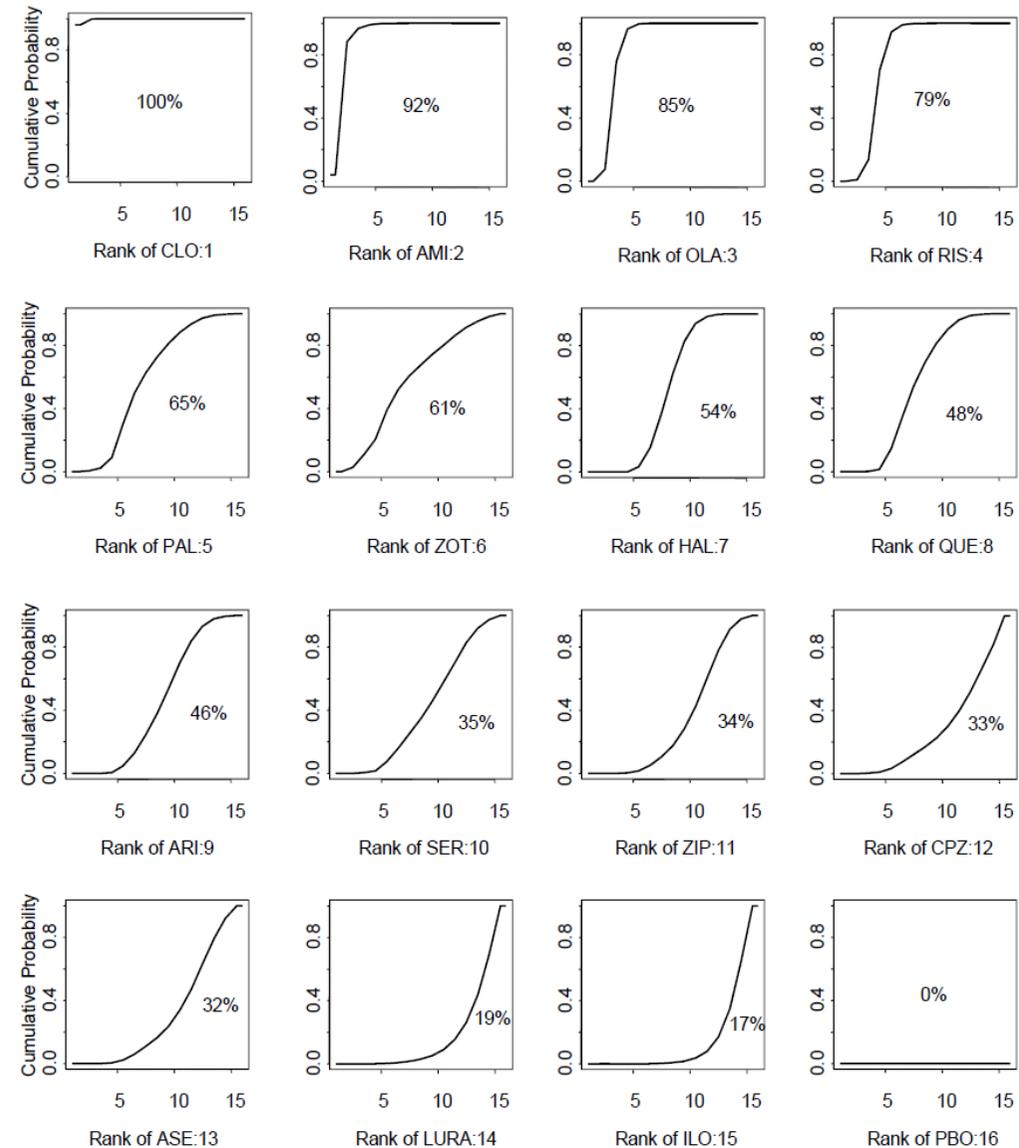
- ランキングプロット
- SUCRA (右図)
 - 線の下側の面積が大きいほど良い順位

• とても分かりやすい指標

- 順位というわかりやすい指標によって優劣をつけられるので馴染みやすい

• この結果だけが先走ってしまう懸念

- 評価のために考慮する情報を度外視している(試験規模・試験数, 推定精度などを完全に反映していない)



今回のNMAの限界

- 古い抗精神病薬

- Perphenazine, sulpirideなどは組み入れられなかった
 - Perphenazineの試験は数少ないが, effectiveness研究において良好な結果を得ている

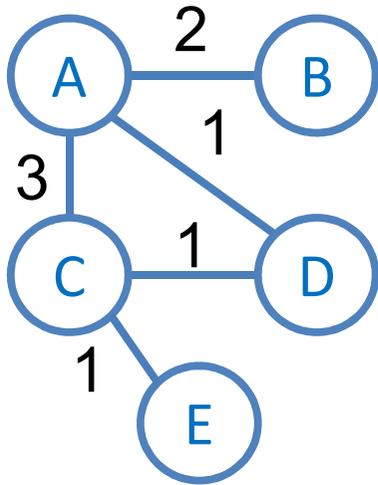
- 一般化の限界

- NMA均一な試験で解析することが求められるので急性期治療, そして投与開始後6週間の反応に限定して論文を抽出した
 - 今回の解析では以下の患者を除外した
 - 若年の患者, 陰性症状が優勢の患者, 再発の患者, 慢性化した患者
 - 短期間の試験に限定した
 - 慢性的な疾患ではあるものの, 一定の期間(6週間)での反応の評価も重要

グループディスカッション1:追加の説明 - 事例(モデル) -

藤井 陽介

論文でのモデリングを理解するために簡略化した仮想NMA



試験 [i]	治療	
	被検 [j]	基準 [k]
1	B	A
2	B	A
3	C	A
4	C	A
5	C	A
6	D	A
7	D	C
8	E	C

- 背景
 - AからEの順に開発された
 - Aは古典的な治療
 - Bは2番目に出てきた治療
 - Cは今の標準的な治療
 - D, Eは比較的最近の治療
- 各治療の開発における検証試験
 - BとCはAと比較した(優越性)
 - DはAと優越性, Cとは非劣性
 - EはCと非劣性
- その他の特徴
 - Cは標準治療なので, D, Eからは対照薬されている
 - ACDはループを成している

記法(論文と同じ記法)

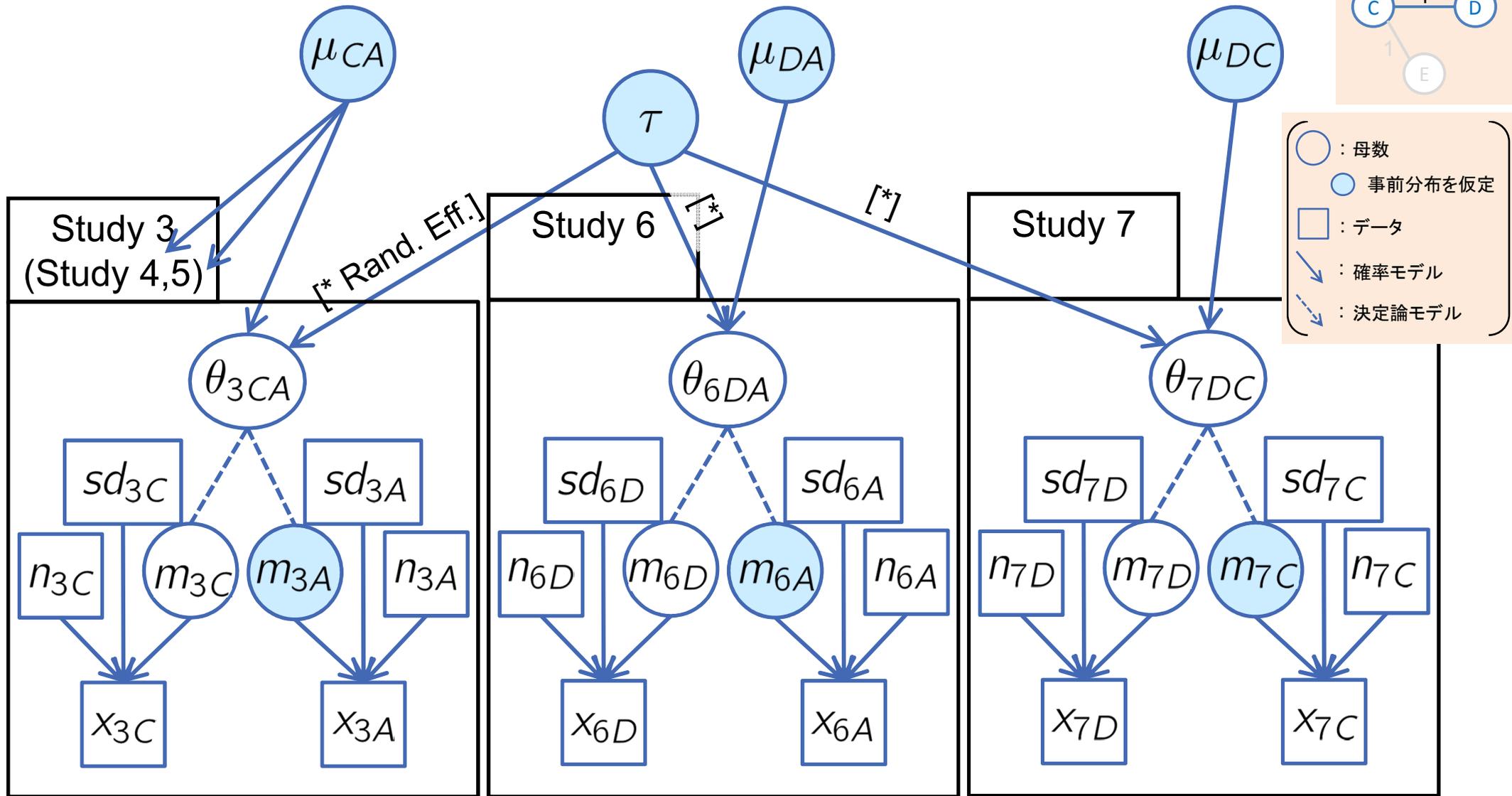
- i : 試験 ($i = 1, \dots, N$) [N個の試験]
- j, k : 治療 ($j, k = A, \dots, T$) [T種類の治療]
 - ただし, ij, ik とするとき, j, k は試験 i で対象としている治療集合の要素となり, 特に k は当該試験の比較における基準治療とする
- n_{ij}, n_{ik} : 標本数
- x_{ij}, x_{ik} : 標本平均
- sd_{ij}, sd_{ik} : 標本標準偏差
- m_{ij}, m_{ik} : 平均に関する母数
- θ_{ijk} : 平均値の差に関する母数(試験ごと)
- μ_{jk} : 平均値の差に関する母数
 - μ_{jA} : 被検治療 j のNMAにおける基準治療(A)に対する平均値の差に関する母数

Inconsistencyモデル

- $x_{ik} \sim N(m_{ik}, sd_{ik}/\sqrt{n_{ik}})$: 標本平均に関する確率モデル
- $m_{ij} = \theta_{ijk} + m_{ik}$: 被検治療群の標本平均に関する決定論モデル
- $m_{ik} \sim N(0, 100)$: 基準治療群の標本平均に関する事前分布
- $\theta_{ijk} \sim N(\mu_{jk}, \tau^2)$: 平均値の差に関する効果変量モデル
 - 3群以上の場合は多変量正規分布とし, 共分散に $\tau^2/2$ を規定
- $\mu_{jj'} \sim N(0, 100^2), j \neq j'$: ハイパーパラメータ(平均)に関する事前分布
- $\tau \sim N(0, 1)/(0,)$: ハイパーパラメータ(標準偏差)に関する事前分布

※ 平均値関連の事前分布についてはnormal vague priorとしか記載がない

InconsistencyモデルのDAG : ACDの比較のみ

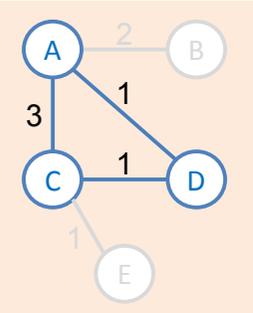
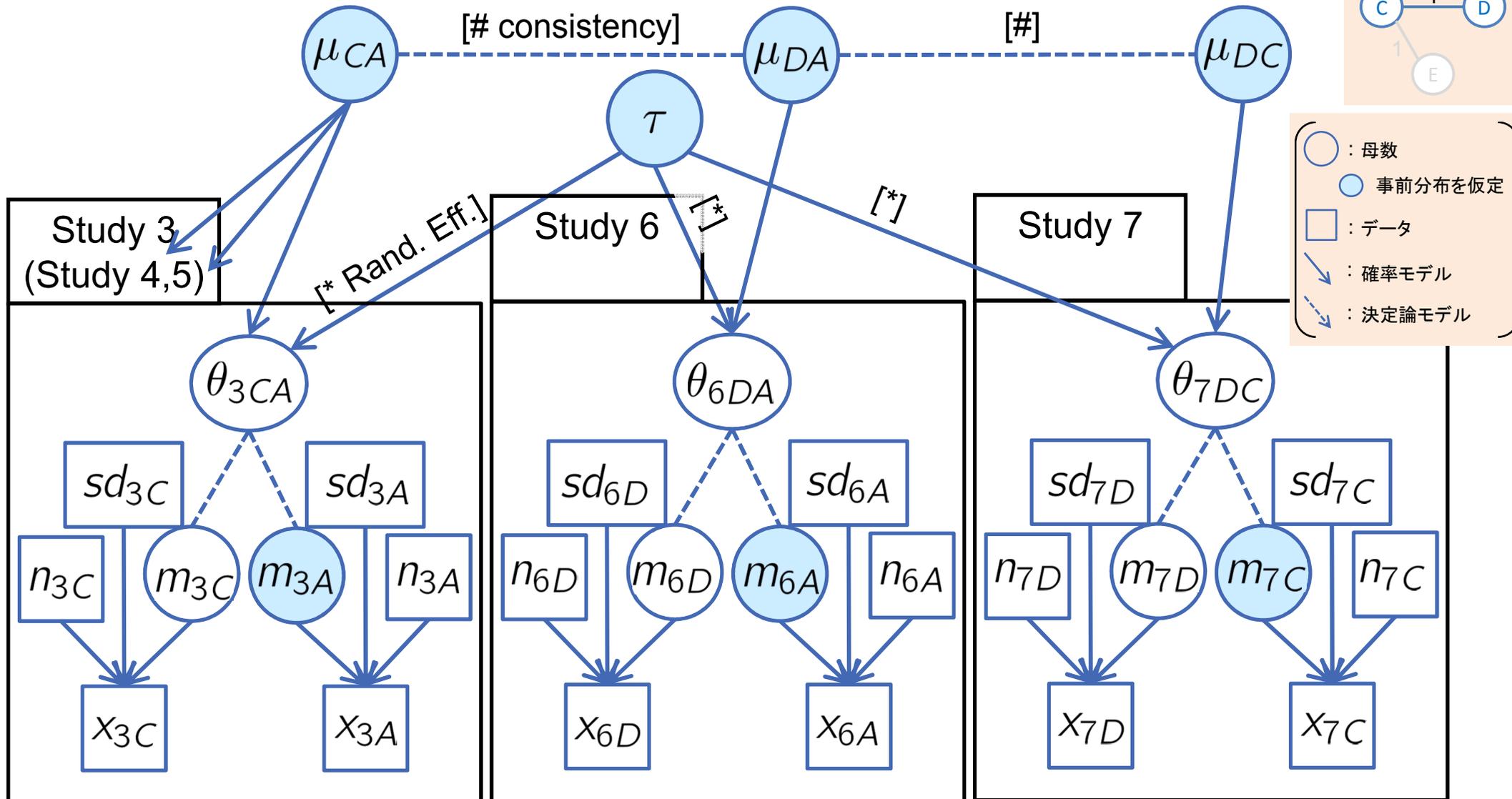


Consistencyモデル

- $x_{ik} \sim N(m_{ik}, sd_{ik}/\sqrt{n_{ik}})$: 標本平均に関する確率モデル
- $m_{ij} = \theta_{ijk} + m_{ik}$: 被検治療群の標本平均に関する決定論モデル
- $m_{ik} \sim N(0, 100)$: 基準治療群の標本平均に関する事前分布
- $\theta_{ijk} \sim N(\mu_{jk}, \tau^2)$: 平均値の差に関する効果変量モデル
 - 3群以上の場合は多変量正規分布とし, 共分散に $\tau^2/2$ を規定
- $\mu_{ja} \sim N(0, 100)$: NMAの基準治療に対する平均値の差に関するハイパーパラメータに関する事前分布
- $\tau \sim N(0, 1) / (0,)$: ハイパーパラメータ(標準偏差)に関する事前分布
- $\mu_{jk} = \mu_{ja} + \mu_{ak} = \mu_{ja} - \mu_{ka}$: 平均値の差の一致性に関する決定論モデル

※ 平均値関連の事前分布についてはnormal vague priorとしか記載がない

ConsistencyモデルのDAG: ACDの比較のみ



- : 母数
- : 事前分布を仮定
- : データ
- : 確率モデル
- - - : 決定論モデル

グループディスカッション1:追加の説明 - ソフトウェア, 情報源 -

兼清 道雄

NMAのためのソフトウェア

- **WinBUGS**

- NICE Decision Support Unit

- **R**: rmeta, netmeta, mvmeta packages

- **R+BUGS系**

- R+WinBUGS, R+OpenBUGS, R+JAGSなど色々

- **Stata**: mvmeta

- セミナー: ネットワーク・メタ・アナリシス入門(丹後)
- ネットワークメタアナリシスの解析モジュール(野間)

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

- **SAS**

- Jones et al. (2011): proc mixed etc.
- Chen (2015): Doing Bayesian Analysis Using PROC MCMC@Deming Conference
- Scott (2014): Network meta-analysis in SAS
- 福井・乙黒・磯崎 (2014) ネットワークメタアナリシスを用いた無作為化比較試験の統合 @SASユーザー総会2014

[http://www.nicedsu.org.uk/Evidence-Synthesis-TSD-series\(2391675\).htm](http://www.nicedsu.org.uk/Evidence-Synthesis-TSD-series(2391675).htm)



NICE Decision Support Unit

NICE National Institute for Health and Care Excellence

- Home
- About the DSU
- Staff
- Appraisal Specific Projects
- Methods Development
- Technical Support Documents
- Training
- Publications

ABOUT THE EVIDENCE SYNTHESIS TSD SERIES

A series of seven TSDs have been produced in the area of Evidence Synthesis. These are:

- TSD 1** [Introduction to evidence synthesis for decision making](#)
- TSD 2** [A general linear modelling framework for pair-wise and network meta-analysis of randomised controlled trials](#) (last updated April 2014)
[WinBUGS system\(.odc\)](#) files (last updated March 2013)
- TSD 3** [Heterogeneity: subgroups, meta-regression, bias and bias-adjustment](#)
[WinBUGS system\(.odc\)](#) files
- TSD 4** [Inconsistency in networks of evidence based on randomised controlled trials](#) (last updated April 2014)
[WinBUGS system\(.odc\)](#) files (last updated March 2013)
- TSD 5** [Evidence synthesis in the baseline natural history model](#)
[WinBUGS system\(.odc\)](#) files

Website Search

Search: Website

Mailing List

Join the DSU mailing list

Name:

Email:

Social networks

Statistical approaches for conducting network meta-analysis in drug development[†]

Byron Jones,^{a*} James Roger,^b Peter W. Lane,^c Andy Lawton,^d Chrissie Fletcher,^e Joseph C. Cappelleri,^f Helen Tate,^g Patrick Moneuse,^h and on behalf of PSI Health Technology Special Interest Group, Evidence Synthesis sub-team

We introduce health technology assessment and evidence synthesis briefly, and then concentrate on the statistical approaches used for conducting network meta-analysis (NMA) in the development and approval of new health technologies. NMA is an extension of standard meta-analysis where indirect as well as direct information is combined and can be seen as similar to the analysis of incomplete-block designs. We illustrate it with an example involving three treatments, using fixed-effects and random-effects models, and using frequentist and Bayesian approaches. As most statisticians in the pharmaceutical industry are familiar with SAS[®] software for analyzing clinical trials, we provide example code for each of the methods we illustrate. One issue that has been overlooked in the literature is the choice of constraints applied to random effects, and we show how this affects the estimates and standard errors and propose a symmetric set of constraints that is equivalent to most current practice. Finally, we discuss the role of statisticians in planning and carrying out NMAs and the strategy for dealing with important issues such as heterogeneity. Copyright © 2011 John Wiley & Sons, Ltd.

Keywords: network meta-analysis; health technology assessment; statistical methodology

オススのメの文獻, Webサイト

• 論文

- 藤井, ネットワークメタアナリシスの医薬品のComparative Effectiveness評価への利用. *薬理と治療*, 2015; **43**: 621-24.
- 野間, メタアナリシスのエビデンスを正しく読み解くために. *薬理と治療*, 2015; **43**: 615-20

• 文献

- 丹後, メタ・アナリシス入門(新版あり)

• Webサイト

- Cochrane Handbook for Systematic Reviews of Interventions
 - <http://handbook.cochrane.org/>
- NICE Decision Support Unit
 - 情報がリッチなTechnical Support Documentがたくさん
 - WinBUGSのプログラムもある
 - [http://www.nicedsu.org.uk/Evidence-Synthesis-TSD-series\(2391675\).htm](http://www.nicedsu.org.uk/Evidence-Synthesis-TSD-series(2391675).htm)
- Cochrane Collaborationのまとめサイト
 - <http://cmim.cochrane.org/methods-innovation-fund-stream-2>

グループディスカッション1:追加の説明 - メタアナリシスの代表的な問題 -

藤井 陽介

メタアナリシスにおける代表的な問題 [野間]

Publication Bias

Mixing Apples and
Oranges

Garbage in, Garbage
out

Publication Bias

- ネガティブな結果が出た臨床試験は公表されにくい
 - 試験実施者にとって都合が悪い結果
 - 論文の査読で評価されにくい
- この状況でメタアナリシスをするとう治療効果を過大評価してしまう
 - 多くのポジティブな結果に基づくメタアナリシスが起こすバイアス
- Turnerらによる報告(右図)
 - FDAの試験登録システムを活用
 - FDAがポジ → 95%公表
 - 疑わしい → 50%
 - ネガ → 33%

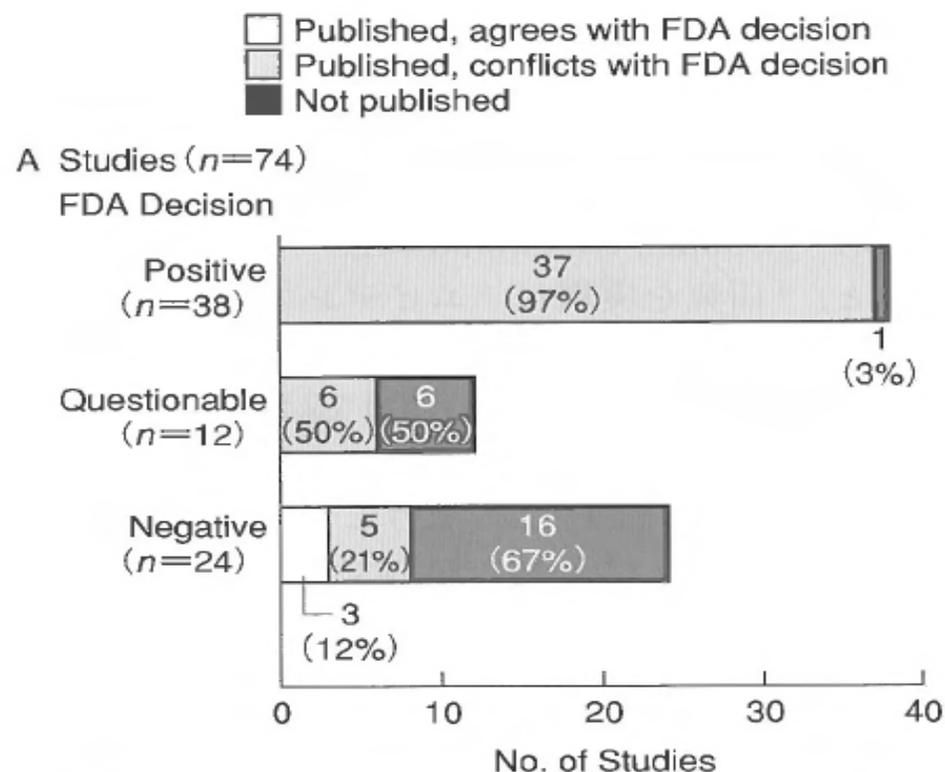


図2 TurnerらによるFDAの臨床試験レジストレーションを利用した出版バイアスの評価
出典 *N Engl J Med* 2008; 358: 252-60. (引用文献6)

[野間]

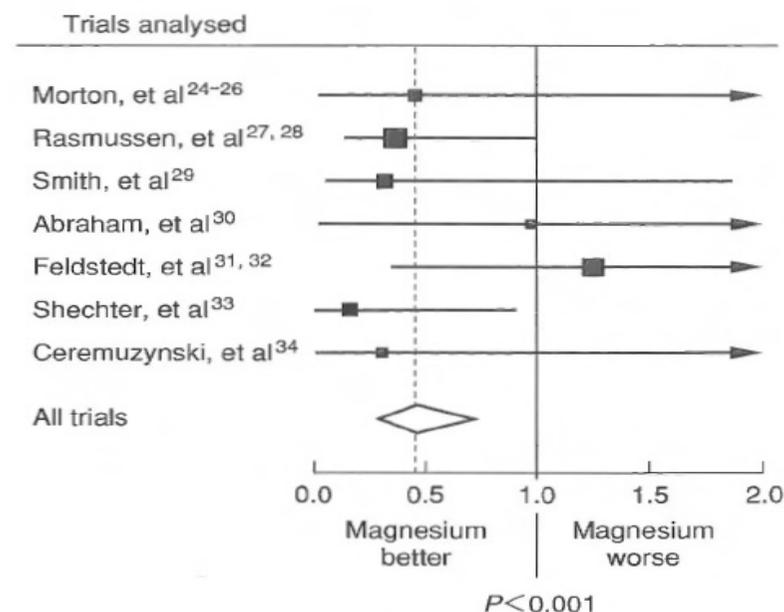
Mixing Apples and Oranges

- メタアナリシスの前提: 各試験の結果が本質的に異質ではない
- しかし, 結果に影響を与えてしまう重要な要因が試験間で異質となってしまうことのほうが一般的
 - 選択・除外基準, 用量・用法, 対照薬の種類, 評価項目の定義
 - 施設の背景(規模など), 医療環境(国・地域間の差)
- メタアナリシスは, 異質なものを強引に統合する危ない方法論?



Garbage in, Garbage out

- そもそも個々の試験の質が低ければ、統合した結果も・・・
- このようなメタアナリシスの結果はバイアスがあるかも
 - 小規模だから質が低いとは限らないかもしれないが



Odds ratios (magnesium: control) for mortality in patients with suspected acute myocardial infarction. ■=Odds ratio and 99% confidence interval (areas of the squares are proportional to amount of information contributed by trial), ◇=95% confidence interval for the overview. Vertical solid line denotes no difference in mortality between treatment and control, while the vertical broken line denotes a 55% reduction in the odds of death

図1 Teoらによるメタアナリシスの結果, 要約オッズ比 0.47 (95% CI: 0.28-0.79)

出典 *BMJ* 1991; 303: 1499-1503. (引用文献2)

[野間]

グループディスカッション1:追加の説明
- システムティックレビュー, エビデンスレベル -

藤井 陽介

NMA論文の全ては「研究を統合して得られた推定値」ではない

- メタアナリシスはシステマティックレビュー(以下, **SR**)の一部にすぎない
 - SRの特性(Cochrane Handbook [Higgins, et al.])
 - 目的および事前に規定された研究の組入基準の明確な提示
 - 厳密で, 再現可能な方法
 - 組入基準を満たす全ての研究を同定することを意図した系統的な検索
 - 組み入れた研究の結論の妥当性の評価。例:risk of biasの評価
 - 組み入れた研究の特性や結論の系統的な提示や**統合**
 - メタアナリシスは「統合」に相当 (メタアナがSRに相当するという考え方もあるが)
 - “Many systematic reviews contain meta-analyses. Meta-analysis is the use of statistical methods to summarize the results of independent studies (Glass 1976).” (by Cochrane Handbook)
- 統合といっても, 単に各試験の推定値を統合することのみならず
 - メタアナリシスは試験間のエビデンスの一致性の検討, および試験間の相違の探索の助けとなる (by Cochrane Handbook)
 - つまり, 統合結果に対するSRの影響を考察することも「統合」に含まれている
 - SRプロセスの各々が信頼できないならメタアナしても…
 - “Mixing Apples and Oranges”や“Garbage in, Garbage out” [野間]

Levels of Evidence 2011 (抜粋, 一部改変) [OCEBM, 2011]

OCEBM: Oxford Centre for Evidence-Based Medicine

アップデートされています

レベル	内容
1*	Systematic review of randomized trials or <i>n</i> -of-1 trials
2*	Randomized trial or observational study with dramatic effect
3*	Non-randomized controlled cohort/follow-up study**
4*	Case-series, case-control studies, or historically controlled studies**
5	Mechanism-based Reasoning

SRが上位



SRは
「エビデンスの王者」
か？



質, 精度, PICO,
試験間の異質性,
推定値の大きさ
について言及し,
それに基づきレベルが
下がることも
組み込まれている

グループディスカッション1:追加の説明 - NMAの活用・実施 -

藤井 陽介

NMAの活用・実施

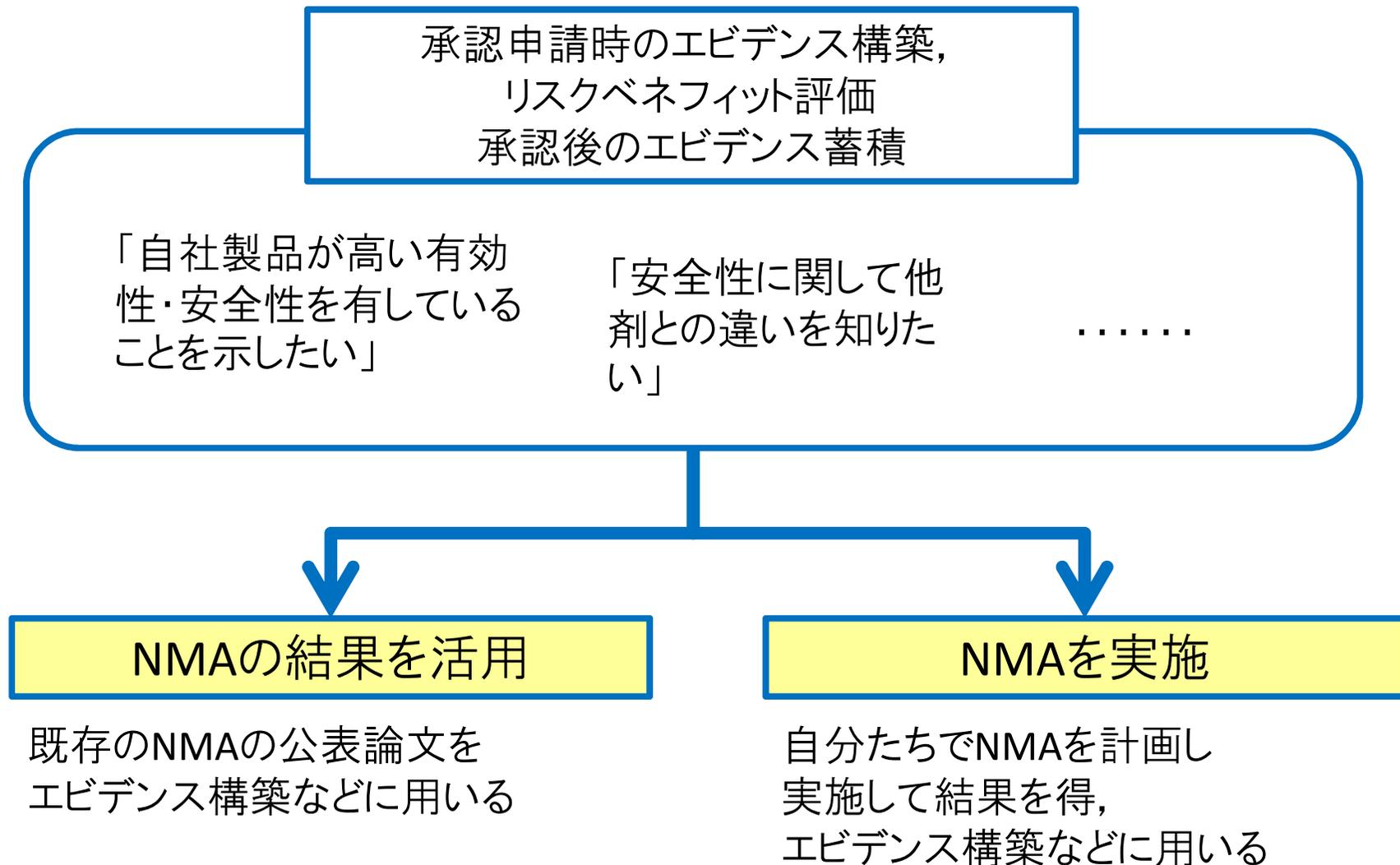
NMAの結果を活用

既存のNMAの公表論文を
効果の評価に用いる

NMAを実施

自分たちでNMAを計画し
実施して結果を得、
効果の評価に用いる

企業におけるNMAの結果の活用・NMAの実施



NMAの結果を活用

- 既存の**NMA**の公表論文を参照する
 - **NMA**の公表論文を引用するときに生じる選択バイアス
 - 「いい結果のようなので, 採用」
 - 「都合が悪そうだから, 不採用」
 - 「微妙なので, 我々にとって良いところだけ採用」
 - 結果だけ追い求めず, 目的と合った**NMA**か吟味が必要
 - 著名な論文誌の論文だから信頼できる? …そうとは限らないかも
 - 結果を活用する者に倫理性・リテラシーが問われる
 - 前述した妥当性・信憑性評価のためのチェックリストが役にたつ

NMAを実施

- 自分たちでNMAを計画し実施する
 - 企業が実施したNMAの結果に対する公表バイアス
 - 「いい結果が出てきたので、公表」
 - 「都合が悪そうなので、お蔵入り」
 - 「都合が悪そうなので、いい部分だけを公表」
 - 査読の段階で様々な観点から客観的評価を受ける
 - ときにやり直し！ときにリジェクト！！
 - 製薬メーカーが関連したNMAは受け入れられるのか？
 - 偏った方法・解釈をしているのではないかという視点で査読されるのは必然
 - 実施・報告する者に倫理性・リテラシーが問われる
 - 報告方法に関する文献は未だ少ない [Bafeta, et al.]
 - 前述のチェックリストを報告者の観点から解釈するのは一案

PROSPERO

Home

- Home
- Register a review
- My PROSPERO records
- My details
- Search PROSPERO**
- Search CRD databases
- About PROSPERO
- Inclusion criteria
- Help with registration
- News
- Support for PROSPERO
- References and resources
- Contact
- Disclaimer

Home > Search PROSPERO

[Sign in or Join](#)

Search

Combine these selections with **AND** ▼

in **All fields** ▼

in **Review title** ▼

in **Review question** ▼

in **Condition/Domain** ▼

in **Participants/Population** ▼

Review status **Any review status** ▼

Date registered to

Search now **Go**

Search by registration number

Search now **Go**

Display all published records **Go**

Search Results [50 records]

Registration no. ⌵	Title ⌵	Status ⌵
CRD42015024930	Gingko biloba for tardive dyskinesia in schizophrenia: a meta-analysis of randomized controlled trials	Ongoing
CRD42015024608	The effects of Tai Chi for schizophrenia: a systematic review and meta-analysis	Ongoing
CRD42015024198	Huperzine A for cognitive deficits in schizophrenia: a systematic review and meta-analysis	Ongoing
CRD42015022129	Tai chi for schizophrenia [Cochrane Protocol]	Ongoing
CRD42015022127	Asenapine versus placebo for schizophrenia [Cochrane Protocol]	Ongoing
CRD42015022024	Haloperidol discontinuation for schizophrenia [Cochrane Protocol]	Ongoing
CRD42015020661	Assessing social networks in patients with schizophrenia: a systematic review of instruments	Ongoing
CRD42015020523	Efficacy and safety of adjunctive aripiprazole in patients with schizophrenia-spectrum disorder: a meta-analysis of randomized controlled trials	Ongoing
CRD42015020455	Variations of metabolic syndrome rates by country-related dietary patterns in people with schizophrenia spectrum and other psychotic disorders	Ongoing

グループディスカッション1:追加の説明 - 他の話題 -

藤井 陽介

NMAの解析方法

- 相対指標を統合する(break randomizationをしない)

- 各試験で推定された相対指標は, 試験間の異質性につながる要因(プラセボ効果, 患者背景などの効果など)を除去できている(=群間で相殺)
- 例: Aを介したBとCの間接比較

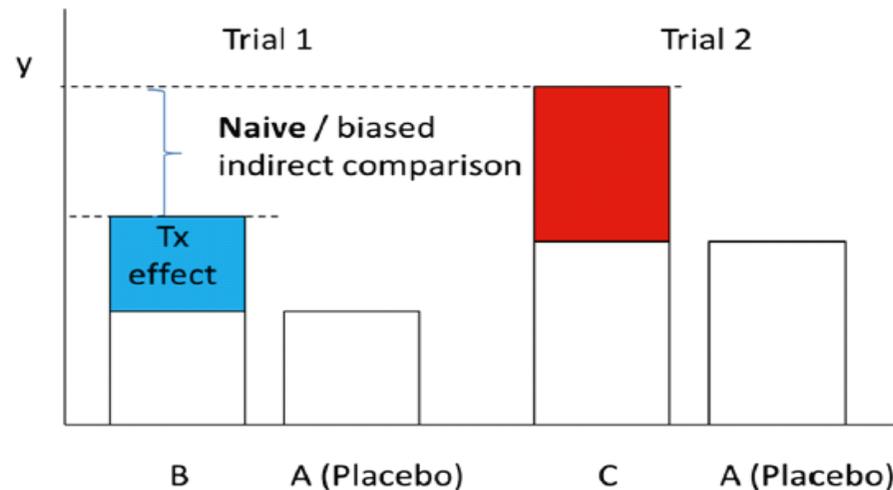
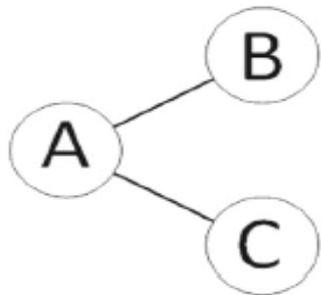


Fig. 9 – Naive indirect comparison that is invalid because differences in study effects are not acknowledged. Note. In this example, the difference in the treatment effect of C versus B is overestimated. Tx, treatment.

[Jansen, et al., 2014]

引用文献(グループディスカッション1)

- Bafeta A, et al. Reporting of results from network meta-analyses: methodological systematic review. *BMJ*, 2014. **348**, g1741.
- Chen FK. Doing Bayesian Analysis Using PROC MCMC. The 71st Deming Conference on Applied Statistics, 2015. (abstract)
- 藤井. ネットワークメタアナリシスの医薬品のComparative Effectiveness評価への利用. *薬理と治療*, 2015. **43**: 621-24.
- Higgins JPT, et al. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. The Cochrane Collaboration, 2011. Available from <http://handbook.cochrane.org/> (アクセス日:2016年1月29日)
- Jansen JP, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. *Value Health*, 2011. **14**, 417-28.
- Jansen JP, et al. Indirect Treatment Comparison/Network Meta-Analysis Study Questionnaire to Assess Relevance and Credibility to Inform Health Care Decision Making: An ISPOR-AMCP-NPC Good Practice Task Force Report. *Value Health*, 2014. **17**, 157-73.
- Jones B, et al. Statistical approaches for conducting network meta-analysis in drug development. *Pharm Stat*, 2011. **10**, 523-31. doi: 10.1002/pst.533.
- Leucht S, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet*, 2013. **382**, 951-62.
- 野間. メタアナリシスのエビデンスを正しく読み解くために. *薬理と治療*, 2015. **43**: 615-20.

引用文献(グループディスカッション1)

- OCEBM Levels of Evidence Working Group. “Oxford Centre for Evidence-based Medicine – Levels of Evidence (March 2009)”. Oxford Centre for Evidence-Based Medicine, 2009. Available from <http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/> (アクセス日:2016年1月29日)
- OCEBM Levels of Evidence Working Group. “The Oxford 2011 Levels of Evidence”. Oxford Centre for Evidence-Based Medicine, 2011. Available from <http://www.cebm.net/index.aspx?o=5653> (アクセス日:2016年1月29日)
- Salanti G, et al. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol*, 2011. **64**: 163–71.
- Scott DA. Network meta-analysis in SAS. Danish Society of Biopharmaceutical Statistics, 2014. Available from http://www.google.co.jp/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=1&cad=rja&uact=8&ved=0ahUK Ewi3yK_4joHLAhWIOJQKHcj-DQEQFgggMAA&url=http%3A%2F%2Fwww.dsbs.dk%2Fmoder%2FDanish%2520SAS%2520course%2520final.pptx&usg=AFQjCNGHpqIBBh3OGeNzi5Tj5doKKAw9QA&sig2=w15yc6E3a5xzdkZROWUvEA. (アクセス日:2016年2月18日)
- 丹後. メタ・アナリシス入門. 朝倉書店, 2016.
- NICE Decision Support Unit. ABOUT THE EVIDENCE SYNTHESIS TSD SERIES. NICE Decision Support Unit. Available from [http://www.nicedsu.org.uk/Evidence-Synthesis-TSD-series\(2391675\).htm](http://www.nicedsu.org.uk/Evidence-Synthesis-TSD-series(2391675).htm). (アクセス日:2016年2月18日)