

メタ分析

専修大学 国里愛彦

2021年3月1日

自己紹介

- 国里愛彦（くにさとよしひこ）
- 所属：専修大学人間科学部心理学科 准教授
- 専門：計算論的アプローチを用いたうつ・不安の維持と認知行動療法の作用メカニズム研究（計算論的臨床心理学）
- 研究の方法論や枠組みを調べるのが好き、オープンサイエンスも好き
- 研究に関することなら気軽にお声掛けください。連絡は、メールフォームからどうぞ！



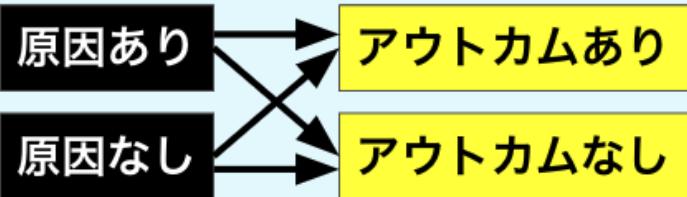
メタ分析（系統的展望）とは？

- 明確な研究疑問に従って、過去の研究を系統的・包括的に収集し、批判的吟味をした上で（系統的展望）、量的に統合する（メタ分析）研究方法
- メタ分析はピアソンから始まり、アイゼンクの心理療法の効果の議論の後に、スミスとグラスによってまとめられた研究法方法になる（山田・井上,2012）。
- 90年代以降の**EBM**の発展に伴って、臨床疫学で実施方法の洗練化が進む。本発表は、臨床疫学でのメタ分析を軸におきつつ、心理学での実施法を解説する。

観察研究

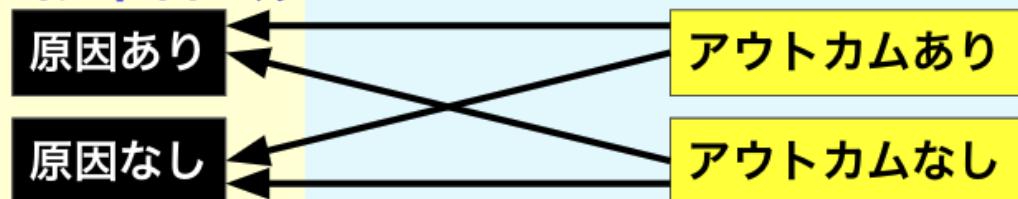
クロスセクショナル

＜疫学の研究デザイン＞

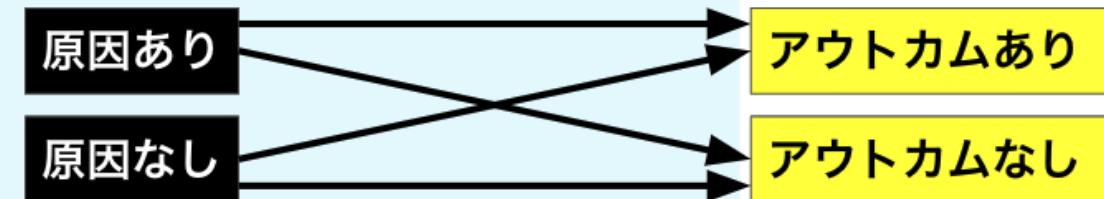


※観察研究での原
因は、疫学では曝
露(Exposure)と呼
ぶ

ケースコントロール



コホート



介入研究

RCT



エビデンスの階層表(Greenhalgh, 1997)

- エビデンスの階層表において、**系統的展望とメタ分析は、最も強いエビデンス**になる。

1. 系統的展望とメタ分析
2. RCT（決定的な差のある）
3. RCT（決定的な差はない）
4. コホート研究
5. ケースコントロール研究
6. クロスセクショナル研究
7. ケース報告
8. 偉い雰囲気の人の意見、動物実験など（国里加筆）



- しかし！適切に実施された系統的展望に限る！

研究報告ガイドライン

臨床研究では、研究報告ガイドラインが整備されている。

1. 系統的展望とメタ分析:**PRISMA声明** (日本語解説は、卓他(2011)や国里(2015))
 2. RCT:**CONSORT声明**(拡張版もあり心理療法など非薬物療法的介入版もある。日本語解説は、奥村(2014))
 3. 観察研究：**STROBE声明**(日本語解説は、竹林(2014))
 4. 健康関連質問紙の評価と作成：**COSMINチェックリスト**(日本語解説は、土屋(2014))
 5. ケース報告:**CAREガイドライン**(日本語解説は、谷(2015))
-
- 以降は、PRISMA声明とコクランの資料作業標準書(Higgins & Thomas, 2019)とCochrane Handbook for Systematic Reviews of Interventions(Higgins et al., 2019)を中心に説明する。

PRISMA声明の拡張版

- PRISMA for Protocols (プロトコル作成)
- PRISMA for Scoping Reviews (スコーピングレビュー, 日本語解説は友利他, 2020)
- PRISMA for Abstracts (学会や雑誌の抄録)
- PRISMA Equity (健康格差)
- PRISMA Harms (有害事象をアウトカムに含むレビュー)
- PRISMA Individual Patient Data (個人レベルの患者データのメタ分析)
- PRISMA for Network Meta-Analyses (ネットワークメタ分析)
- PRISMA for Diagnostic Test Accuracy (診断精度研究)
- PRISMA for Acupuncture (鍼治療)
- PRISMA for Searching (検索)

APAのMeta-Analysis Reporting Standards (MARS)

- APAもPRISMAなどを参考にメタ分析の報告ガイドラインを作成している (Cooper, H., 2020)。しかし、PRISMAと大枠では変わらないし、期待した心理学特有の質の評価などの情報は少ないので、PRISMAでいいかも？

レビュー論文の種類

レビューは、大まかに以下の3つに分けられる(卓他,2020)。

1. **システムティックレビュー（系統的展望）**：プロトコルの事前登録、透明性の高い検索方略とデータ抽出、個々の研究の質の評価、メタ分析と総合的な提案を含めた考察を満たすレビュー
2. **スコーピングレビュー**：プロトコルの作成、透明性の高い検索方略とデータ抽出を満たすレビュー（システムティックレビューの前段階に位置づけられる。発見的なレビュー）
3. **ナラティブレビュー**：上記を満たしていないレビュー
 - ・システムティックレビュー→スコーピングレビュー→ナラティブレビューの順で求められるものが多くなる。

系統的展望の手順

1. 研究疑問を作成し、スコープを決める
2. 研究の検索
3. 研究の選択とデータ収集
4. バイアスのリスクの評価
5. メタ分析
6. 結果の評価と解釈

系統的展望の手順

1. 研究疑問を作成し、スコープを決める
2. 研究の検索
3. 研究の選択とデータ収集
4. バイアスのリスクの評価
5. メタ分析
6. 結果の評価と解釈

研究疑問を作成し、スコープを決める

- 臨床研究の場合、**PICO**（患者、介入、比較対象、アウトカム）や**PECO**（患者、曝露、比較対象、アウトカム）を使って疑問を定式化する。

→心理学においても、どのような対象に対して、どのような独立変数が、どのような従属変数に影響を与えるか明確にする。

- **アウトカム（心理学だと従属変数）を事前に定める**（臨床研究だと有益なアウトカムと有害なアウトカムを1つずつ選ぶ、最大7つ）
- アウトカムの測定方法は複数あるので（例.うつ症状を測定する多数の尺度），レビューに含めるものを事前に定める。

適格基準を事前に定める

- 適格基準を事前に決めておかないと、研究者の求める結果に合わせて研究の組入・除外をしてしまうかも。
- 対象者，介入・曝露（独立変数），対照条件，アウトカム（従属変数），研究デザイン，言語・地域，出版状況（未発表データをいれるか）を明確に決めておく。
- 研究の多様性を包含できるように十分に広く設定するといが、意味のある結論が導ける程度に狭くしておく必要がある（精神障害への心理療法の効果は広すぎる、神奈川県の私大で4月に国里式瞑想法を実施した際のGW後の出席回数への影響は狭すぎる）。

事前登録

- 系統的展望は、**事前にプロトコルを作つてから実施する必要がある。**
- 研究疑問・適格基準を決め、その後のレビューの方法（検索、バイアスのリスク評価、結果の統合方法、サブグループ解析の計画、結果の評価方法）をプロトコルに記載し、事前登録サイトに登録する。
- **PROSPERO:**系統的展望の事前登録サイト（登録までに時間がかかる、レビューの種類によっては却下される）
- **osf.io:** オープンサイエンスフレームワークのリポジトリ(事前登録はすぐに可能であり、心理学ではおすすめ)

系統的展望の手順

1. 研究疑問を作成し、スコープを決める
2. **研究の検索**
3. 研究の選択とデータ収集
4. バイアスのリスクの評価
5. メタ分析
6. 結果の評価と解釈

研究の検索

出版バイアスを小さくするために、**複数の方法を用いて、包括的・系統的に文献を探す**

1. 一般的な書誌データベースの検索 (CENTRAL, MEDLINE, Embaseなど)
2. 専門書誌データベースの検索(テーマに合わせて、国別、地域別、テーマ別の書誌データベースを検索する。心理系だとPyscINFOなど)
3. 臨床試験登録サイトを検索(Clinical Trials, OSFのRegistrationsなど)
4. 灰色文献の検索(報告書、学位論文、卒論、学会抄録などの未公表データ)
5. 過去の同じトピックのレビュー内検索
6. 入手した文献や関連するレビューの引用文献リストの検索
7. 未公表研究の関係者や団体へ問い合わせをする

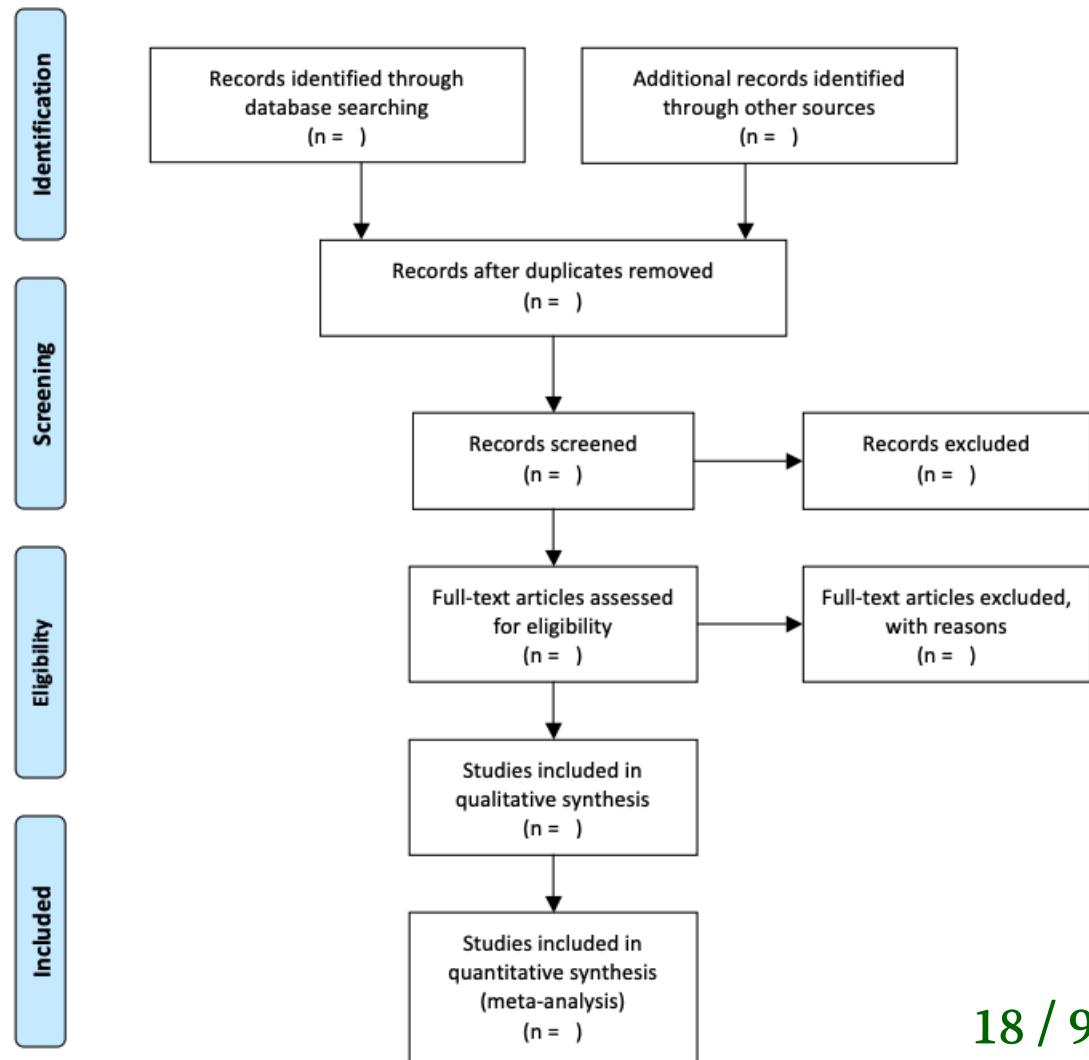
検索方略を構築する

- **検索式**を作る（基本はPとIと研究デザインを指定する。心理学の場合は、IやEとOかもしれない。感度を最大にして、妥当な精度が担保される検索式を検討する。ANDやOR演算子を適切に利用する）
- **統制用語**の使用(MEDLINEのMeSHは医学関係の用語集で論文と関連づけられており検索で有用), **フリーテキスト用語の工夫**(異なったつづり, 同義語, 標準頭字語, 前方一致, 近接演算子など), **検索フィルター**の使用 (検索データベースごとに設定されている)
- 検索が再現できるように, **検索過程（情報源, 時期, 担当者, 検索用語や検索式）について文章化**しておく。

系統的展望の手順

1. 研究疑問を作成し、スコープを決める
2. 研究の検索
3. 研究の選択とデータ収集
4. バイアスのリスクの評価
5. メタ分析
6. 結果の評価と解釈

PRISMA Flow diagram 検索して特定した論文数, スクリーニングで残った論文数, 適格基準を満たした論文数, メタ分析に含んだ論文数を記録する



研究のスクリーニングと適格基準の検討

- (少なくとも) **2人の人間が独立に作業し**, 意見の相違があった場合に解決する方法を決める。
- 最初にタイトルと要約に基づいたスクリーニングを実施し, 残った研究に対し全文に基づいた適格性を判断する。
- **適格基準を満たした研究は, 使用可能なデータがないとしても, レビューに含める** (利用可能なデータが報告されてないこともわかるようにしておく)
- スクリーニングは除外した研究の数の報告だけでよいが, **適格基準から除外した研究は, 理由が必要になる**

具体的な研究選択手順

1. 書誌DBなどで文献検索して、文献情報をエクスポート
2. 文献管理ソフト（Zoteroなど）にインポートして統合
3. 文献管理ソフトで重複を削除する
4. 文献管理ソフトからRIS形式でエクスポートし、レビュー支援ソフトのRayyanにインポートする
5. Rayyan上で、**2名が独立して**タイトルとアブストでスクリーニング(不一致は話し合ったり、第3者を入れて解決)
6. 研究の全文を入手し、同じ研究の複数の報告はまとめる
7. Rayyan上で、**2名が独立して**全文で適格基準を満たすか検討
(不一致は話し合ったり、第3者を入れて解決。不明点は原著者に問い合わせ。Rayyanは全文アップロード不可なので、クラウドストレージを活用)

Rayyan QCRI

- Rayyan QCRIは、スクリーニングや適格性の検討において「**2人の人間が独立に**作業し、意見の相違があった場合に解決する」のを支援してくれるウェブアプリケーション。
- 文献管理ソフトからタイトル・アブストをインポートでき、組入・除外の評価がサクサクできる。
- 独立にブラインドで評価ができる仕組みが用意されている（意見が一致しない論文もまとめてくれるので楽）。



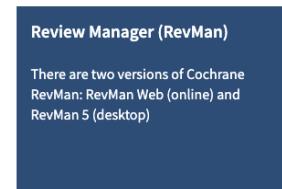
Rayyan QCRI

研究からデータを収集する

- 各研究の文献から**研究の特性**(例えば、参加者、介入、比較条件、アウトカム、研究デザイン)と**アウトカムデータ**を抽出するため、(少なくとも)**2人の人間が独立に作業し**、意見の相違があった場合に解決する方法を決めておく。
 - Cochrane共同計画のデータ抽出フォームがあるが、RCT用なので、心理学や研究テーマに合わせて調整する必要がある。
- データ抽出方法：(1)紙、(2)Googleフォーム等、(3)RevMan等のソフト
 - 各種データ抽出ツールをまとめたサイトが参考になる。上記以外は有料が多い。RevManは便利だが、心理学用にカスタマイズする場合は、Google Formの方がやりやすいかも

RevMan

- Cochrane共同計画が作成するメタ分析用のソフトウェア
- プロトコルの作成, 研究選択, データ抽出, 分析, 報告書の作成までできる。
- 高機能なソフトウェアだが(報告されている数値から逆算して, メタ分析で使う数値を計算する機能もある), コクランレビュー向けの機能も多いので, コクランレビューでないなら, データ抽出やメタ分析で使うくらいかも?



データ抽出時の注意

- 研究の**正誤表や撤回文**がないか確認する（研究発表後に見つかった問題点によっては、評価やデータ抽出に影響がある可能性がある）
- **未出版データ**を入手する（論文に記載されてない情報やまだ論文になつてない情報がある場合は、できるだけ著者に問い合わせをして入手することが望ましい）
- データ抽出者間で結果が一致するように、**事前にデータ抽出フォームの使い勝手などを検討**する（論文では直接報告されてないので計算して求める場合の方法なども事前に決めておく）
- データ抽出は手動で行われるのでミスが混入することを前提にして、**入力ミスのチェック**をする（例えば、フォレストプロットが元論文と同じ方向に同じくらいの効果を示している確認する）

系統的展望の手順

1. 研究疑問を作成し、スコープを決める
2. 研究の検索
3. 研究の選択とデータ収集
4. バイアスのリスクの評価
5. メタ分析
6. 結果の評価と解釈

研究の質の評価

- スミスとグラスのメタ分析に対して、アイゼンクが「garbage in, garbage out」と批判（山田・井上, 2012）。
→グラスとスミスは、質の評価をメタ分析に組み込んで批判に反論した
- 心理学でのメタ分析において、個々の研究の質の評価は難しいとされるが（山田・井上, 2012），疫学領域では研究デザインに従った個々の研究の評価ツールが開発されている。

Risk of Biasの評価

- (少なくとも) **2人の人間が独立に**個々の研究の**バイアスのリスクを評価**し、意見の相違があった場合に解決する方法を決める。その根拠も記載する。
- コクランではRCTの評価にRoB2, 非ランダム化介入研究の評価にROBINS-Iを使用するとされるが([参考リンク](#)), 介入研究以外のツールも整備が進んでいる。
 - RoB2:ランダム化比較試験の質の評価
 - ROBINS-I:ランダム化していない介入研究の質の評価
 - QUADAS2:診断精度研究の質の評価
 - COSMIN Risk of Bias tool:健康関連尺度の質の評価
 - ROBINS-E :観察研究の質の評価(開発中)
- 今後は心理学でも必要?

RoB2: 個々の研究について、5ドメイン内の質問に回答した上で、バイアスリスクについて、アルゴリズムに従ってLow, Some concerns, Highの3段階評価

1. **ランダム化過程におけるバイアス**(割付の順序がランダムではない、割付の隠蔽化ができるない、ベースラインに差がある等)
 2. **意図した介入からの逸脱によるバイアス**(参加者やスタッフの盲検化ができない、盲検化しておらず意図した介入からの逸脱がある、ITTなどの適切な解析が行われていない等)
 3. **アウトカムの欠測によるバイアス**(アウトカムの欠測がある、欠測がありバイアスがある、欠測が介入効果に依存する等)
 4. **アウトカムの測定によるバイアス**(アウトカムの測定法が不適切、測定に群間差がある、評価者が介入情報を知っている等)
 5. **選択的報告バイアス**(アウトカムが複数の方法で測定・分析されており、それが事前登録で計画されていない)
- **総体評価**：個々の研究をLow(全ドメインがLow), Some concerns(Highはないけど、1つ以上のドメインでconcernはある), High(1つ以上のドメインでHighがあるor複数のドメインでconcernsがあり)で評価

RoB2の例

| | Risk of bias domains | | | | | |
|-----------------|----------------------|----|----|----|----|---------|
| | D1 | D2 | D3 | D4 | D5 | Overall |
| Toyo (1928) | + | + | + | + | - | - |
| Kinki (1943) | X | X | X | X | + | X |
| Isumei (1922) | + | + | + | + | + | + |
| Keio (1920) | + | + | + | - | - | - |
| Angaku (1920) | - | + | - | - | - | X |
| Shishsha (1924) | X | X | - | X | X | X |
| Tokai (1924) | - | + | + | + | + | - |
| Asakura (1924) | + | + | + | + | + | + |

Domains:

- D1: Bias due to randomisation.
- D2: Bias due to deviations from intended intervention.
- D3: Bias due to missing data.
- D4: Bias due to outcome measurement.
- D5: Bias due to selection of reported result.

Judgement

- X High
- Some concerns
- + Low

バイアス評価の利用

- 評価したバイアスのリスクはメタ分析に組み込む
 - 1. 感度分析に使う（バイアスの高い研究を入れる時と入れない時を比較する）
 - 2. バイアスリスクに応じて層別解析
 - 3. バイアスリスクが高い研究を除外して解析する
- エビデンスの確実性（独立変数と従属変数の因果関係についての確信度）の評価に、バイアスリスクの総体評価を組み込む(GRADE アプローチ)。

系統的展望の手順

1. 研究疑問を作成し、スコープを決める
2. 研究の検索
3. 研究の選択とデータ収集
4. バイアスのリスクの評価
5. メタ分析
6. 結果の評価と解釈

メタ分析の一般的手順

1. 個々の研究の効果量を計算
2. 個々の研究の効果量を統合
3. 統計学的異質性を評価
4. 事前に計画したサブグループ解析
5. 感度分析
6. 非報告バイアス（出版バイアス）の検討

メタ分析の一般的手順

1. 個々の研究の効果量を計算
2. 個々の研究の効果量を統合
3. 統計学的異質性を評価
4. 事前に計画したサブグループ解析
5. 感度分析
6. 非報告バイアス（出版バイアス）の検討

連続変数における2群の差の効果量

2群の差は、標準化平均値差(Hedgesのg)を用いる

- $Hedges\text{の}g = \frac{\text{介入群の平均} - \text{統制群の平均}}{\text{2群をプールした標準偏差}}$
- 2群をプールした標準偏差 = $\sqrt{\frac{(n_{\text{介入}} - 1)\hat{\sigma}_{\text{介入}}^2 + (n_{\text{統制}} - 1)\hat{\sigma}_{\text{統制}}^2}{(n_{\text{介入}} + n_{\text{統制}} - 2)}}$

* $\hat{\sigma}^2$ は不偏分散。他にCohenのdやGlassの Δ もあるが、バイアスを修正した効果量のHedgesのgがよく使用される

- 標準化平均値差の分散 = $\frac{n_{\text{介入}} + n_{\text{統制}}}{n_{\text{介入}} n_{\text{統制}}} + \frac{g^2}{2(n_{\text{介入}} + n_{\text{統制}})}$
- 標準化平均値差の標準誤差 = $\sqrt{\text{標準化平均値差の分散}}$

分散と標準偏差は効果量とサンプルサイズで決まる

2つの連續変数の関連の効果量

2変数間の関連の強さは、ピアソンの積率相関係数を用いる

- $r = \frac{\text{共分散}_{x,y}}{\text{標準偏差}_x \text{ 標準偏差}_y}$
- 相関係数の分散 = $\frac{(1-r^2)^2}{n-1}$
- 統計学的特性の望ましさから、フィッシャーのZ変換を用いることが多い。
- $z = 0.5 \ln\left(\frac{1+r}{1-r}\right)$
- z の分散 = $\frac{1}{n-3}$

2値変数（あり・なし）の効果量

1. リスク差=介入群のありの比率-統制群のありの比率

2. 治療必要数 = $\frac{1}{\text{リスク差}}$

3. リスク比 = $\frac{\text{介入群のありの比率}}{\text{統制群のありの比率}}$

4. オッズ比 = $\frac{\frac{\text{介入群のありの比率}}{1-\text{介入群のありの比率}}}{\frac{\text{統制群のありの比率}}{1-\text{統制群のありの比率}}}$

リスク差や治療必要数(Number needed to treat)は解釈しやすいが、計算上オッズ比が便利なので、オッズが使用される。さらに対数オッズ比($\ln(\text{オッズ比})$)を使う（1を起点に解釈→0を起点に解釈、分散も計算がしやすい）。

オッズ比の分散 = 介入群ありの比率 + 介入群なしの比率 + 統制群ありの比率 + 統制群なしの比率

メタ分析の一般的手順

1. 個々の研究の効果量を計算
2. **個々の研究の効果量を統合**
3. 統計学的異質性を評価
4. 事前に計画したサブグループ解析
5. 感度分析
6. 非報告バイアス（出版バイアス）の検討

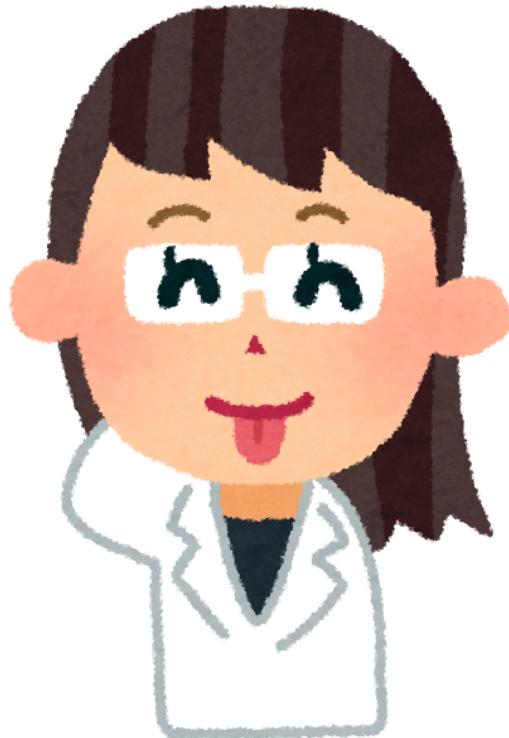
メタ分析は難しくない。

それぞれの研究の効果量を重みをつけて
平均するだけ！

え、 それだけ？

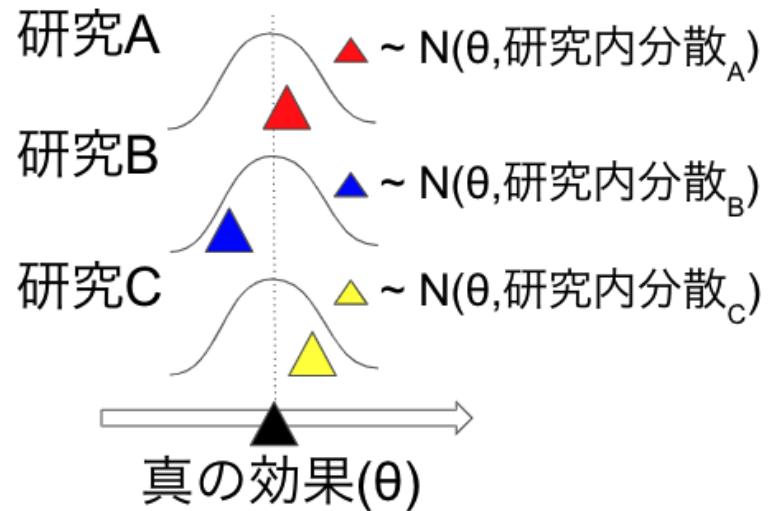


それだけ～ モデルや推定法でチョット変わるけどね...



固定効果モデル

- 1つの真の効果を仮定し、個々の研究の効果は真の効果と研究内分散(標本誤差)によるとするモデル
- General variance based法による平均値差の統合の場合、重みは付けは、各研究の効果量の分散の逆数を使う。



$$\text{統合された効果量} = \frac{\sum W_i Y_i}{\sum W_i} = \frac{(\text{各研究の重み} W_i * \text{各研究の効果量} Y_i) \text{の合計}}{\text{各研究の重み} W_i \text{の合計}}$$

$$W_i = \frac{1}{V_i} = \frac{1}{\text{効果量の分散} V_i}$$

固定効果モデル

- 統合された効果量の標準誤差は、各研究の重みの合計から計算される

$$\text{統合された効果量の標準誤差} = \sqrt{\frac{1}{\text{各研究の重み } W_i \text{ の合計}}}$$

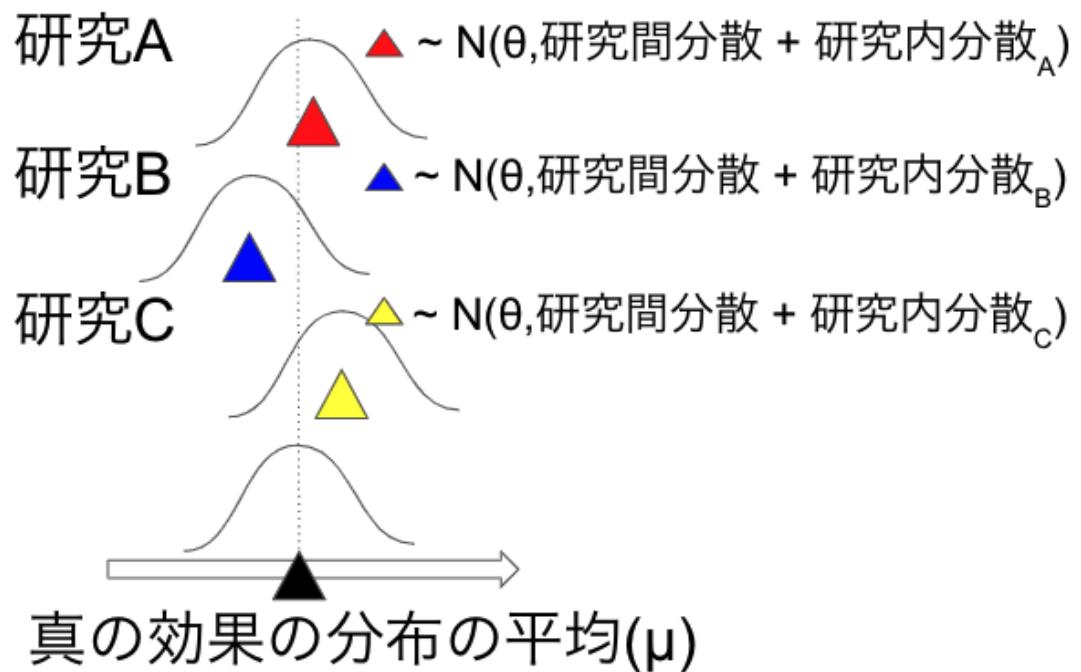
- 95%信頼区間(CI)は、以下のように計算する。

95%CIの下限 = 統合された効果量 - 1.96 * 統合された効果量の標準誤差

95%CIの上限 = 統合された効果量 + 1.96 * 統合された効果量の標準誤差

ランダム効果モデル

- 真の効果を分布として考えて、研究ごとに真の効果が異なると仮定する。
- 個々の研究の効果は、真の効果の分布の平均値(μ)と研究間分散(τ^2)と研究内分散(標本誤差)によるとするモデル



ランダム効果モデル

- 平均値差の統合(DerSimonian-Laird法)では、以下のように、重み付けに研究間分散(τ^2)と研究内分散(V_i)を用いる。

1. 固定効果モデルを用いて統合された効果量を算出し、 Q を計算する(T は個々の研究の効果量 \bar{T} は統合された効果量)。 Q は研究間のばらつきの程度。

$$Q = \sum W_i (\bar{T}_i - \bar{\bar{T}})^2$$

2. Q を用いて、研究間分散(τ^2)を計算する。

$$\tau^2 = \frac{Q - (k-1)}{\sum W_i - \frac{\sum \bar{T}_i}{\sum W_i}}$$

3. 重み付けの計算に $\tau^2 V_i$ を用いる以外は、以降は固定効果モデルと同じ。

$$W_i = \frac{1}{V_i + \tau^2}$$

ランダム効果モデル

- DerSimonian-Lairdの他に, Paule-Mandel, Restricted Maximum-Likelihood, Maximum-likelihood, Hunter-Schmidt, Sidik-Jonkman, Hedges, Empirical Bayesがある(Rのmeta パッケージで利用可能)。
- DerSimonian-Lairdがよく使われるが, それはRevManで使われているなどの理由であって, Maximum-Likelihood, Sidik-Jonkman, Empirical Bayesの方が研究間変動の推定には良い性質がある(Harrer rt al., 2019)
- ランダム効果か固定効果かは事前登録の段階で決める。

メタ分析の一般的手順

1. 個々の研究の効果量を計算
2. 個々の研究の効果量を統合
3. **統計学的異質性を評価**
4. 事前に計画したサブグループ解析
5. 感度分析
6. 非報告バイアス（出版バイアス）の検討

概念的異質性

- 論文を精査することで明らかとなる質的な研究間の違い
- 臨床的異質性：年齢，性別，人種，重症度，合併症など患者集団の違い（心理学だと対象者属性の異質性）
- 方法論的異質性：治療法，治療期間，治療者の経験など方法の違い（心理学だと測定法，実験操作などの違い）

統計学的異質性

- 統計学的に明らかとなる研究ごとの効果のばらつき
- 統計学的異質性も概念的異質性もある場合：(1)メタ分析しない,(2)サブグループ解析・メタ回帰分析の実施

統計学的異質性

- ランダム効果モデルのQを使った検定もあるが、研究数が多いと有意になりやすい（有意じゃなくても検定力低いだけかも）

→ 検定ではなく、 Higgins et al.(2003)の $I^2 = \frac{Q-(k-1)}{Q} 100$ のような記述的指標を用いて判断をする。 I^2 は、効果量のバラツキの内、効果量の異質性が占める割合を%で表している。

- I^2 のおおまかな目安
 - 40%以下：異質性は問題ではないかも
 - 30~60%：中程度の異質性があるかも
 - 50~90%：実際に異質性があるかも
 - 75~100%：無視できない異質性がある

メタ分析の一般的手順

1. 個々の研究の効果量を計算
2. 個々の研究の効果量を統合
3. 統計学的異質性を評価
4. 事前に計画したサブグループ解析
5. 感度分析
6. 非報告バイアス（出版バイアス）の検討

サブグループ解析

- ・異質性を生むと考えられるサブグループを特定して（患者の特性、方法など），分けて効果量の統合をする。

→プロトコルの段階で計画を立てる

感度分析

- ・適格基準や統合方法を変えても今回の結果が頑健に得られるか確認する。

→(1)適格基準の変更（患者の特徴を変えても変わらない？）,(2)分析データの変更（調整済み効果量では？）,(3)分析法の変更(固定効果とランダム効果モデルで大きな差はない？)

メタ分析の一般的手順

1. 個々の研究の効果量を計算
2. 個々の研究の効果量を統合
3. 統計学的異質性を評価
4. 事前に計画したサブグループ解析
5. 感度分析
6. **非報告バイアス（出版バイアス）の検討**

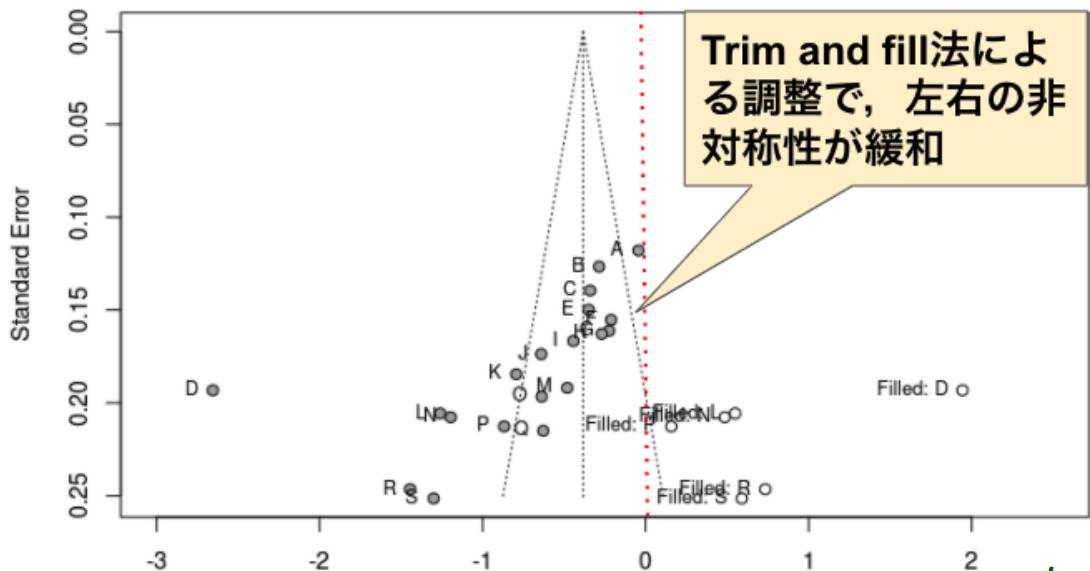
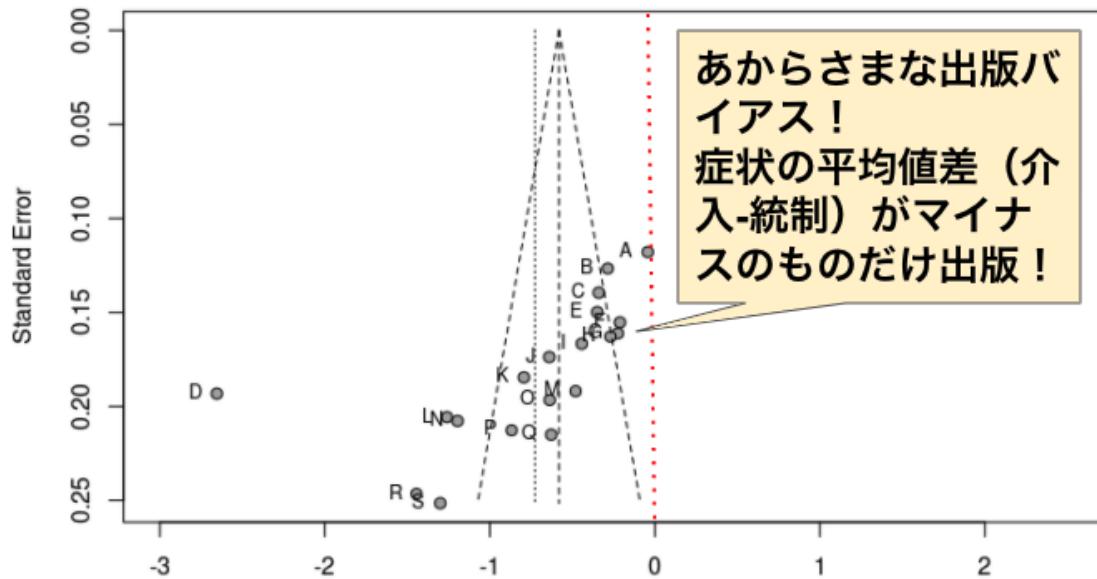
出版バイアス

ネガティブな結果が出版されにくく、効果を過剰評価するかもしれない

(Funnel plotやEggerの検定などで出版バイアスを確認)。

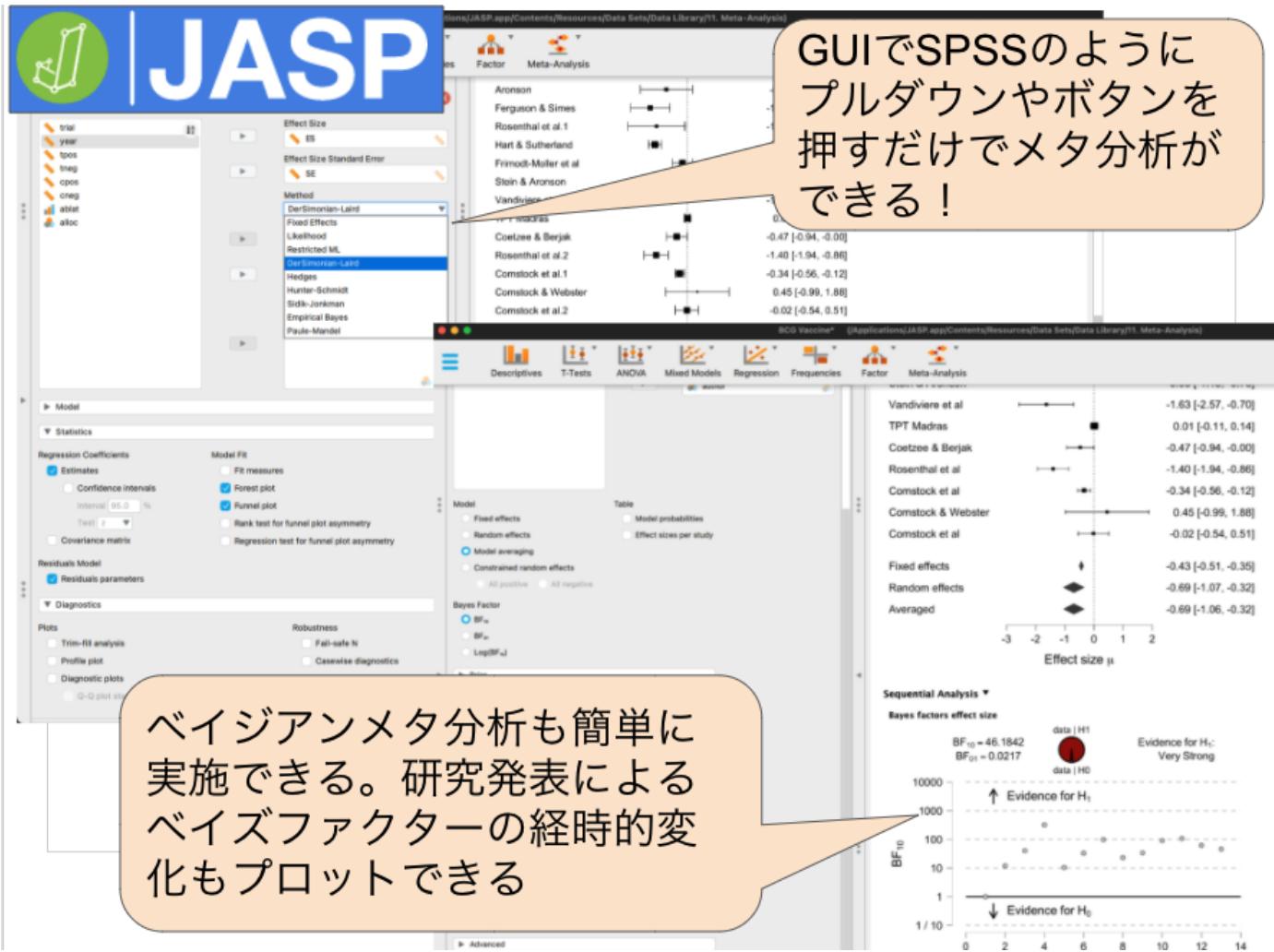
出版バイアスがある場合は、Trim-fill法などで効果の過大評価を調整する。

Trim-fill法:偏った側からデータを除外して、左右が対称になるように調整する。その後、除外したデータを戻して、その反対側にデータを追加する



メタ分析してみよう！

JASP(無料)



STATA(有料)岡田・小野寺(2018)『実践的メタ分析入門』に詳しい

Rでメタ分析

- 個人的にRが好きなので(このスライドもRで作っています), 以降ではRでメタ分析を実施する方法を説明します。
- Rは, GUIじゃないので敬遠される部分もありますが, 逆に言うと, 以降のコードをコピペすれば解析できます。
- Rでのメタ分析の資料としては, 『Doing Meta-Analysis in R : A Hands-on Guide』 (Harrer, M., Cuijpers, P., Furukawa,T.A., Ebert, D.D.)を超える資料はないのではないかと思います。無料で読めます。

Rを使う準備

1.RとRStudioをインストールする

- R
- RStudio

2.Rパッケージをインストールする(使用するのはmetaとdmetarになる。devtoolsはGitHub経由でdmetarをインストールするために必要)

```
install.packages("meta")
install.packages("devtools")
devtools::install_github("MathiasHarrer/dmetar")
```

2群の平均値差:データ準備

```
# パッケージの読み込み
```

```
library(dmetar)
```

```
library(meta)
```

```
# データを作る (Mは平均, Sは標準偏差, Nは人数, eが介入・実験群, cが統制群)
```

```
Me <- c(35, 40, 35, 22, 38, 31, 29, 28)
```

```
Se <- c(7.5, 8.2, 5.9, 10.4, 8.8, 9.1, 6.4, 4.6)
```

```
Ne <- c(15, 32, 18, 22, 27, 42, 29, 17)
```

```
Mc <- c(20, 23, 30, 10, 22, 25, 19, 20)
```

```
Sc <- c(6.5, 4.3, 5.7, 7.8, 6.9, 10.4, 7.1, 6.5)
```

```
Nc <- c(16, 30, 19, 21, 29, 40, 32, 18)
```

```
Author <- c("Senshu (1923)", "Waseda (1920)", "Doshisha (1920)", "Kangaku (1932)", "Keio (1920)", "
```

```
Subgroup <- c("East", "East", "West", "West", "East", "West", "West", "East")
```

```
# データフレームを作る
```

```
data_meta <- data.frame(Author, Me, Se, Ne, Mc, Sc, Nc, Subgroup)
```

2群の平均値差：メタ分析の実施

- metaパッケージの連續変数用メタ分析関数のmetacontを使う。
- comb.fixedが固定効果、comb.randomがランダム効果になる。
- method.smdは、デフォルトが"Hedges"だが、"Cohen"や"Glass"も選べる。
- method.tauは、デフォルトがDL(DerSimonian-Laird)だが他にも選択できる。今回は、最尤法(ML)を選択した。

```
# メタ分析
result_meta <- metacont(Ne,Me,Se,Nc,Mc,Sc,
                         data = data_meta,
                         studlab = paste(Author),
                         comb.fixed = TRUE,
                         comb.random = TRUE,
                         prediction = FALSE,
                         sm = "SMD",
                         method.smd = "Hedges",
                         method.tau = "ML")
```

2群の平均値差：メタ分析結果

```
> result_meta
```

| | SMD | 95%-CI | %W(fixed) | %W(random) |
|-----------------|--------|------------------|-----------|------------|
| Senshu (1923) | 2.0868 | [1.1894; 2.9841] | 6.1 | 10.1 |
| Waseda (1920) | 2.5400 | [1.8606; 3.2194] | 10.7 | 12.3 |
| Doshisha (1920) | 0.8438 | [0.1678; 1.5197] | 10.8 | 12.4 |
| Kangaku (1932) | 1.2770 | [0.6154; 1.9387] | 11.3 | 12.5 |
| Keio (1920) | 2.0041 | [1.3538; 2.6544] | 11.7 | 12.6 |
| Ritsumei (1922) | 0.6093 | [0.1658; 1.0527] | 25.2 | 14.9 |
| Kinki (1943) | 1.4568 | [0.8876; 2.0259] | 15.3 | 13.5 |
| Toyo (1928) | 1.3813 | [0.6347; 2.1279] | 8.9 | 11.6 |

Number of studies combined: k = 8

| | SMD | 95%-CI | z | p-value |
|----------------------|--------|------------------|-------|----------|
| Fixed effect model | 1.3690 | [1.1466; 1.5914] | 12.06 | < 0.0001 |
| Random effects model | 1.4895 | [1.0552; 1.9239] | 6.72 | < 0.0001 |

Quantifying heterogeneity:

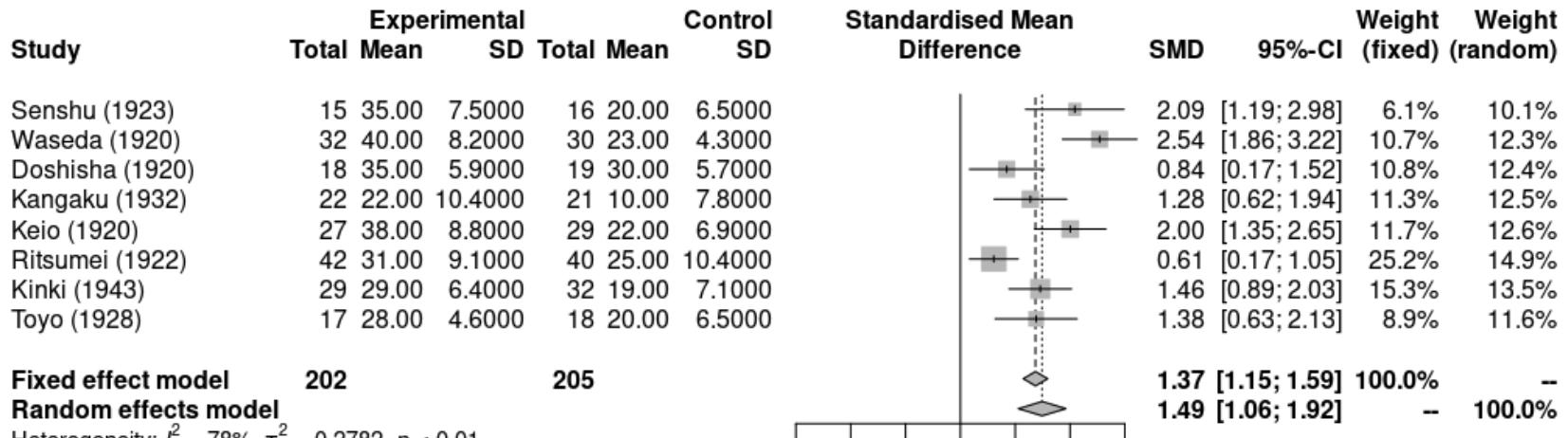
$\tau^2 = 0.2782$ [0.0838; 1.6462]; $\tau = 0.5275$ [0.2895; 1.2830]
 $I^2 = 77.6\%$ [55.8%; 88.7%]; $H = 2.11$ [1.50; 2.97]

Test of heterogeneity:

| Q | d.f. | p-value |
|-------|------|----------|
| 31.30 | 7 | < 0.0001 |

2群の平均値差：フォレストプロット

```
forest(result_meta)
```



2群の平均値差：異質性の確認

```
summary(result_meta)
```

```
> summary(result_meta)
Number of studies combined: k = 8

           SMD      95%-CI     z   p-value
Fixed effect model 1.3690 [1.1466; 1.5914] 12.06 < 0.0001
Random effects model 1.4895 [1.0552; 1.9239] 6.72 < 0.0001

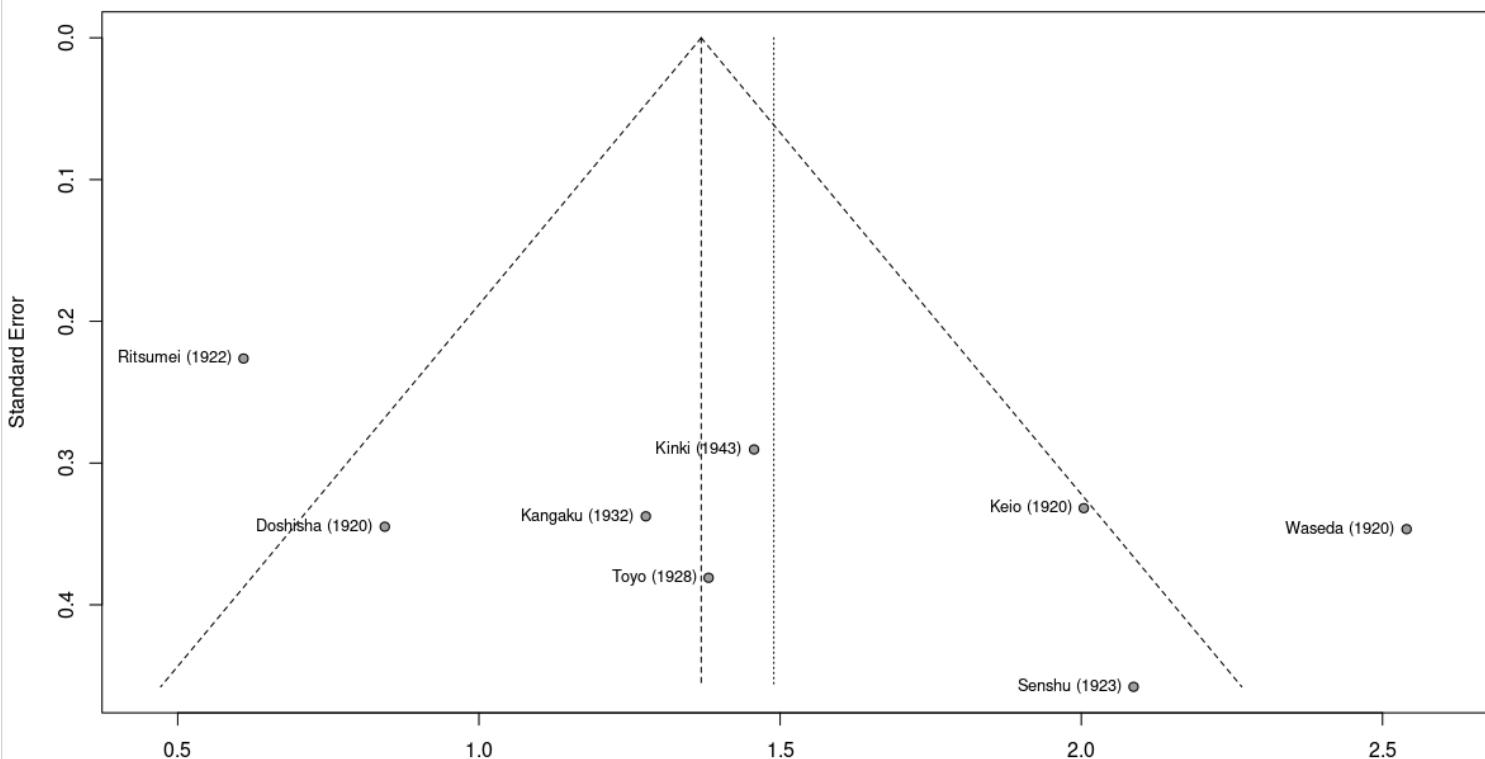
Quantifying heterogeneity:
tau^2 = 0.2782 [0.0838; 1.6462]; tau = 0.5275 [0.2895; 1.2830]
I^2 = 77.6% [55.8%; 88.7%]; H = 2.11 [1.50; 2.97]
```

```
Test of heterogeneity:
  Q d.f.  p-value
31.30    7 < 0.0001
```

- 異質性に関連した外れ値の検討、Influence Analyses、GOSH Plot Analysisは省略しました。詳細は、[Doing Meta-Analysis in R : A Hands-on Guide](#)を確認ください

2群の平均値差：出版バイアス

```
funnel(result_meta, xlab = "g", studlab = TRUE)
```



2群の平均値差：出版バイアス(dmetarパッケージのeggers.test)

```
# Egger's test  
eggers.test(x = result_meta)
```

```
## Warning in eggers.test(x = result_meta): Your meta-analysis contains k = 8  
## studies. Egger's test may lack the statistical power to detect bias when the  
## number of studies is small (i.e., k<10).
```

```
## Eggers' test of the intercept  
## =====  
##  
##   intercept    95% CI      t      p  
##   6.559 0.49 - 12.63 2.118 0.07849926  
##  
## Eggers' test does not indicate the presence of funnel plot asymmetry.
```

2群の平均値差：出版バイアスの調整(trim and fill法)

```
trimfill(result_meta)
```

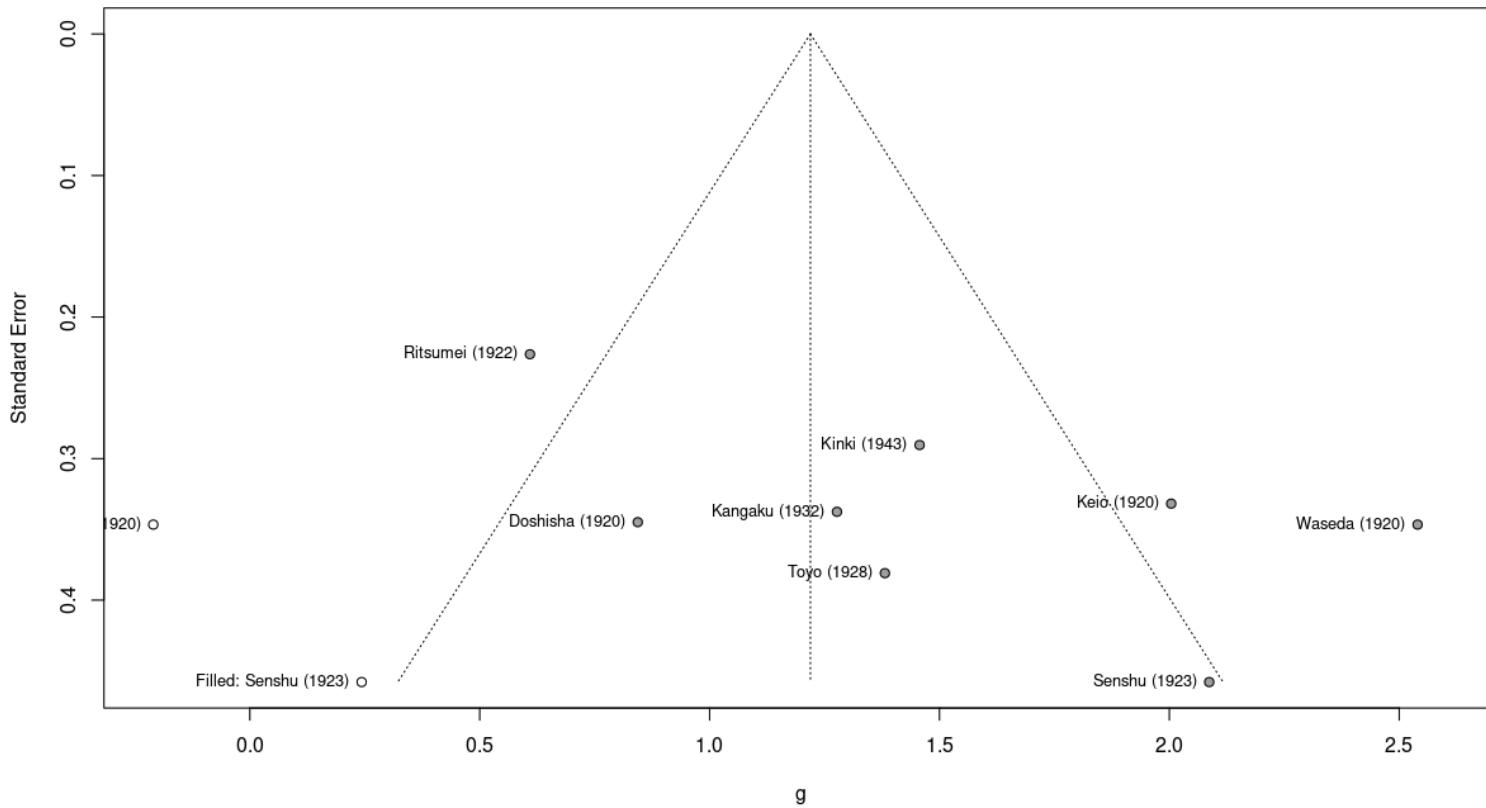
```
> trimfill(result_meta)
      SMD      95%-CI %W(random)
Senshu (1923) 2.0868 [ 1.1894; 2.9841]     8.9
Waseda (1920) 2.5400 [ 1.8606; 3.2194]    10.1
Doshisha (1920) 0.8438 [ 0.1678; 1.5197]    10.1
Kangaku (1932) 1.2770 [ 0.6154; 1.9387]    10.2
Keio (1920) 2.0041 [ 1.3538; 2.6544]    10.2
Ritsumei (1922) 0.6093 [ 0.1658; 1.0527]   11.3
Kinki (1943) 1.4568 [ 0.8876; 2.0259]    10.7
Toyo (1928) 1.3813 [ 0.6347; 2.1279]    9.7
Filled: Senshu (1923) 0.2432 [-0.6542; 1.1405] 8.9
Filled: Waseda (1920) -0.2101 [-0.8894; 0.4693] 10.1
```

Number of studies combined: k = 10 (with 2 added studies)

| | SMD | 95%-CI | z | p-value |
|----------------------|--------|------------------|------|----------|
| Random effects model | 1.2193 | [0.7182; 1.7203] | 4.77 | < 0.0001 |

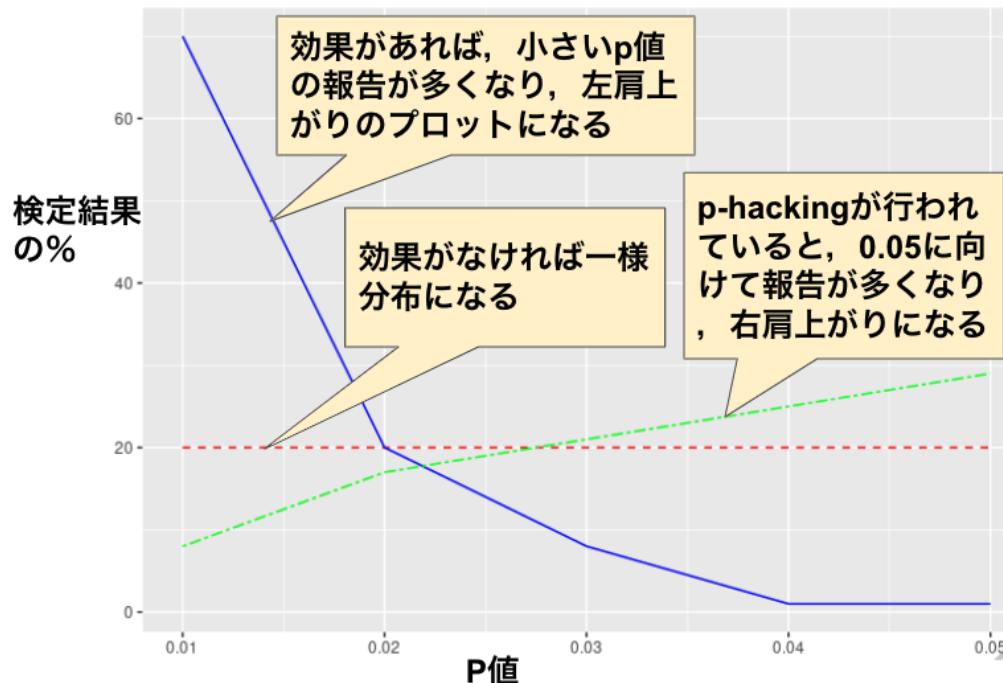
2群の平均値差：出版バイアスの調整結果

```
funnel(trimfill(result_meta),xlab = "g",studlab = TRUE)
```



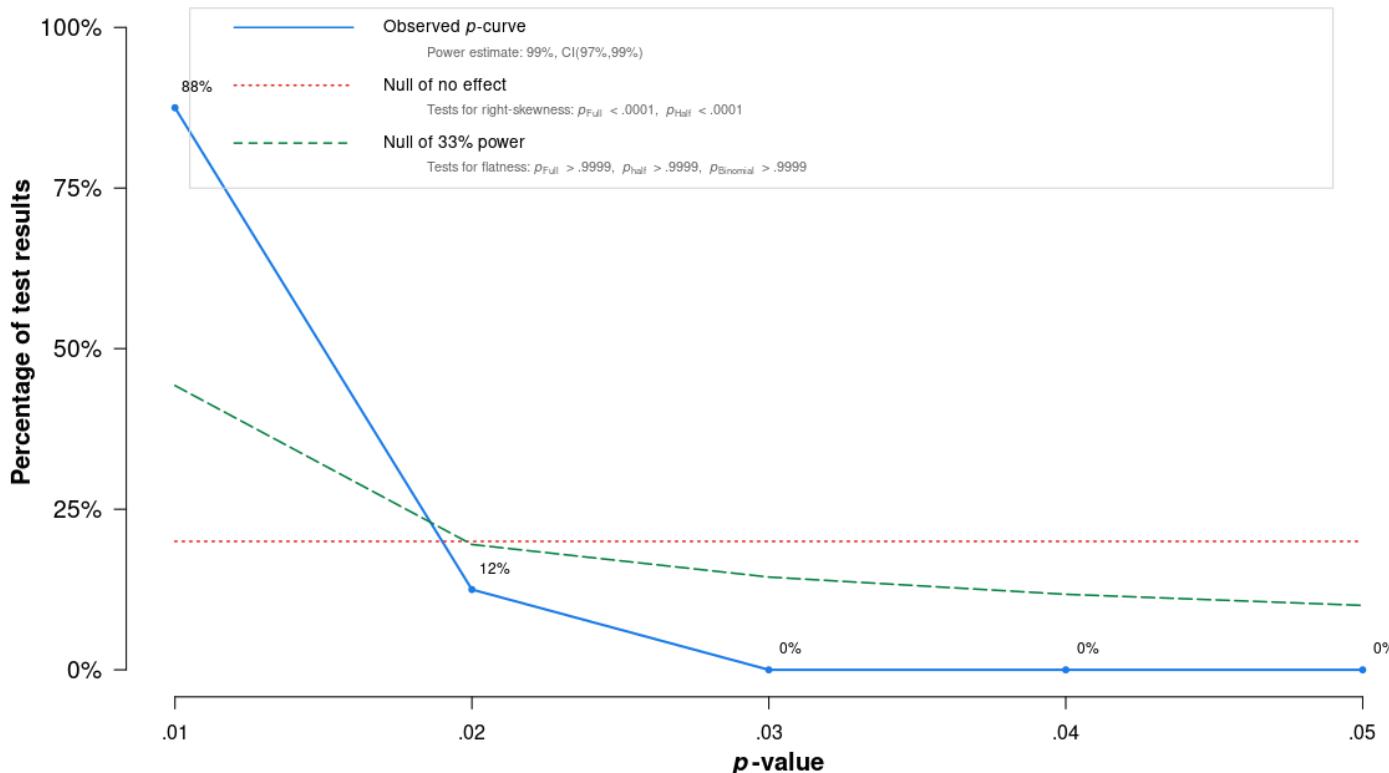
2群の平均値差：P-Curve

- ・サンプルサイズが小さい研究は正確さに欠けるので、それを元にしたTrim and fill法は不正確な効果量を推定するかも。
- ・P-curve分析(Simonsohn et al., 2014)は、p値の報告数から出版バイアスを検討し、真の推定値を推定する方法



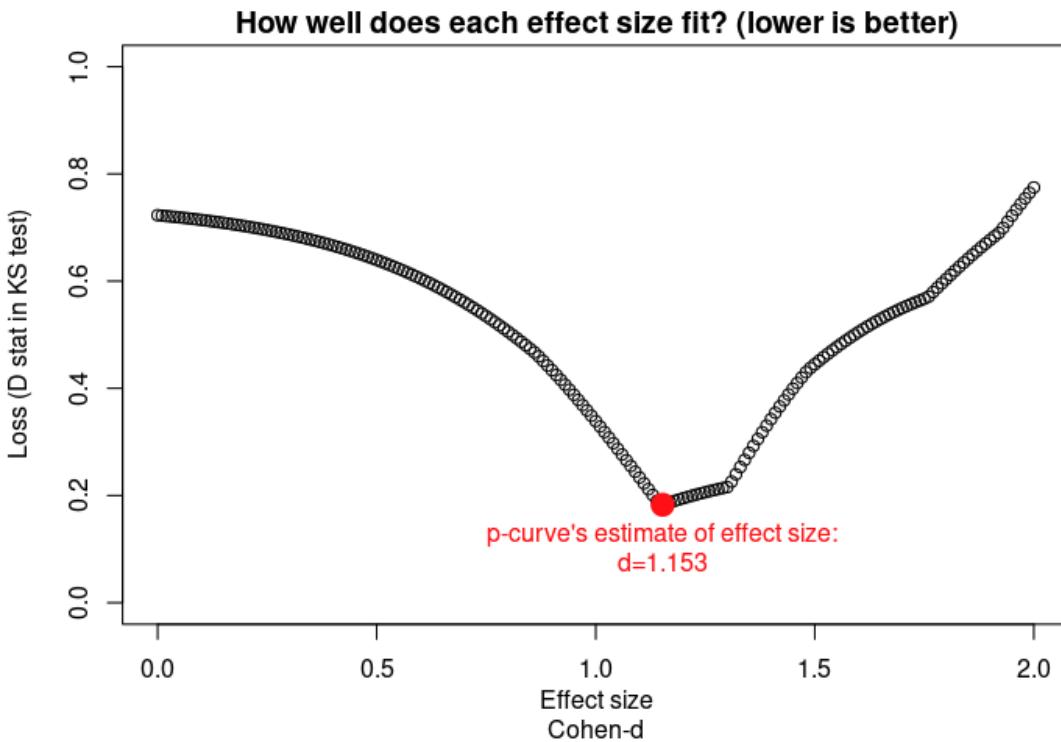
2群の平均値差：P-Curve(dmetar)パッケージのpcurve)

pcurve(result_meta)



2群の平均値差：P-Curve(真の効果の推定)

```
total_sample_size <- Ne + Nc  
pcurve(result_meta, effect.estimation = TRUE, N = total_sample_size, dmin = 0, dmax = 2)
```



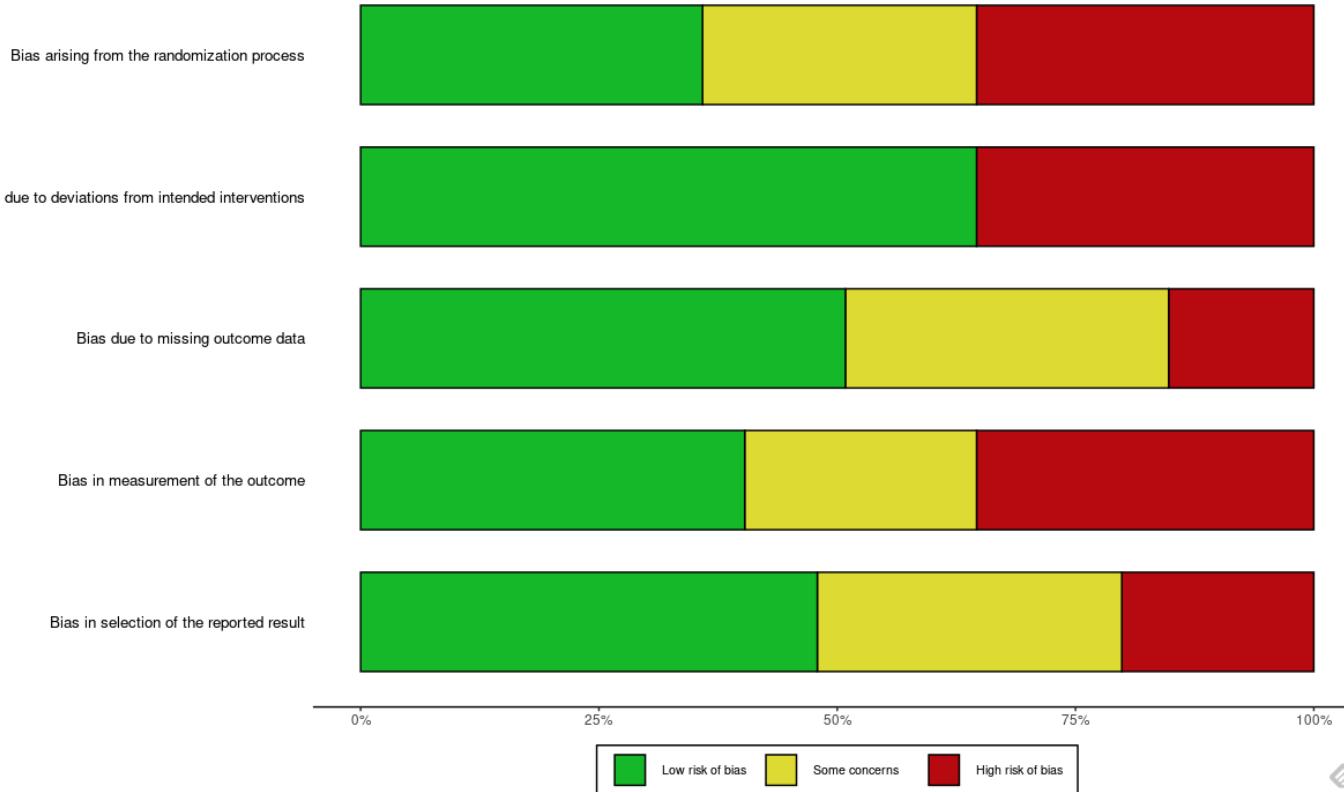
2群の平均値差：RoBプロット(robvisパッケージ)

```
library(robvis)
# RoBの評価をデータフレーム化
# (実際は、エクセルか他のソフトで整理してcsvが良いです)
D1 <- c("Low", "High", "Low", "Low", "Some concerns", "High", "Some concerns", "Low")
D2 <- c("Low", "High", "Low", "Low", "Low", "High", "Low", "Low")
D3 <- c("Low", "High", "Low", "Low", "Some concerns", "Some concerns", "Low", "Low")
D4 <- c("Low", "High", "Low", "Some concerns", "Some concerns", "High", "Low", "Low")
D5 <- c("Some concerns", "Low", "Low", "Some concerns", "Some concerns", "Hihg", "Low", "Low")
Overall <- c("Some concerns", "High", "Low", "Some concerns", "High", "High", "Some concerns",
total_sample_size <- Ne + Nc
data_rob <- data.frame(Study = Author, D1, D2, D3, D4, D5, Overall, Weight = total_sample_size)
head(data_rob)
```

| ## | Study | D1 | D2 | D3 | D4 | D5 |
|------|-----------------|---------------|------|---------------|---------------|---------------|
| ## 1 | Senshu (1923) | Low | Low | Low | Low | Some concerns |
| ## 2 | Waseda (1920) | High | High | High | High | Low |
| ## 3 | Doshisha (1920) | Low | Low | Low | Low | Low |
| ## 4 | Kangaku (1932) | Low | Low | Low | Some concerns | Some concerns |
| ## 5 | Keio (1920) | Some concerns | Low | Some concerns | Some concerns | Some concerns |

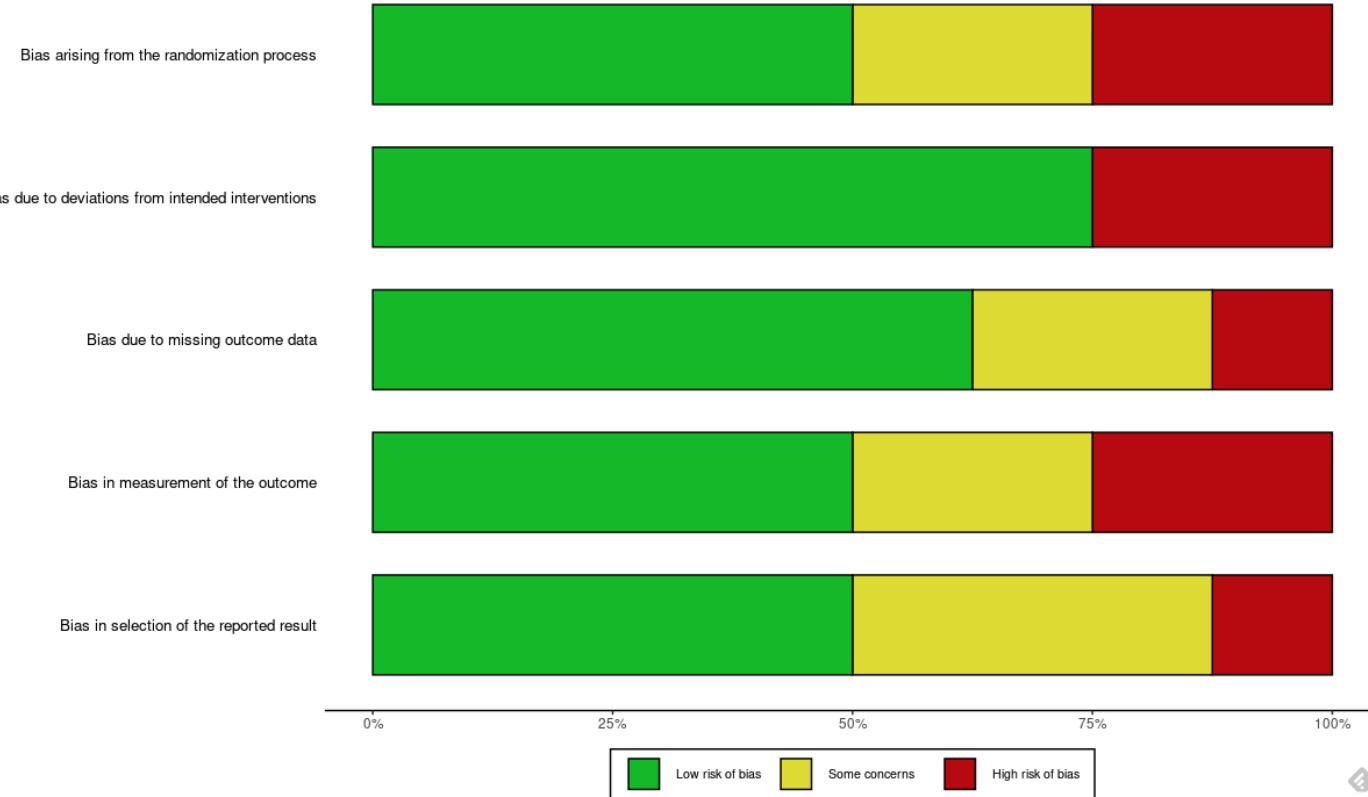
2群の平均値差：RoBプロット(ウェイトあり)

```
rob_summary(data = data_rob, tool = "ROB2")
```



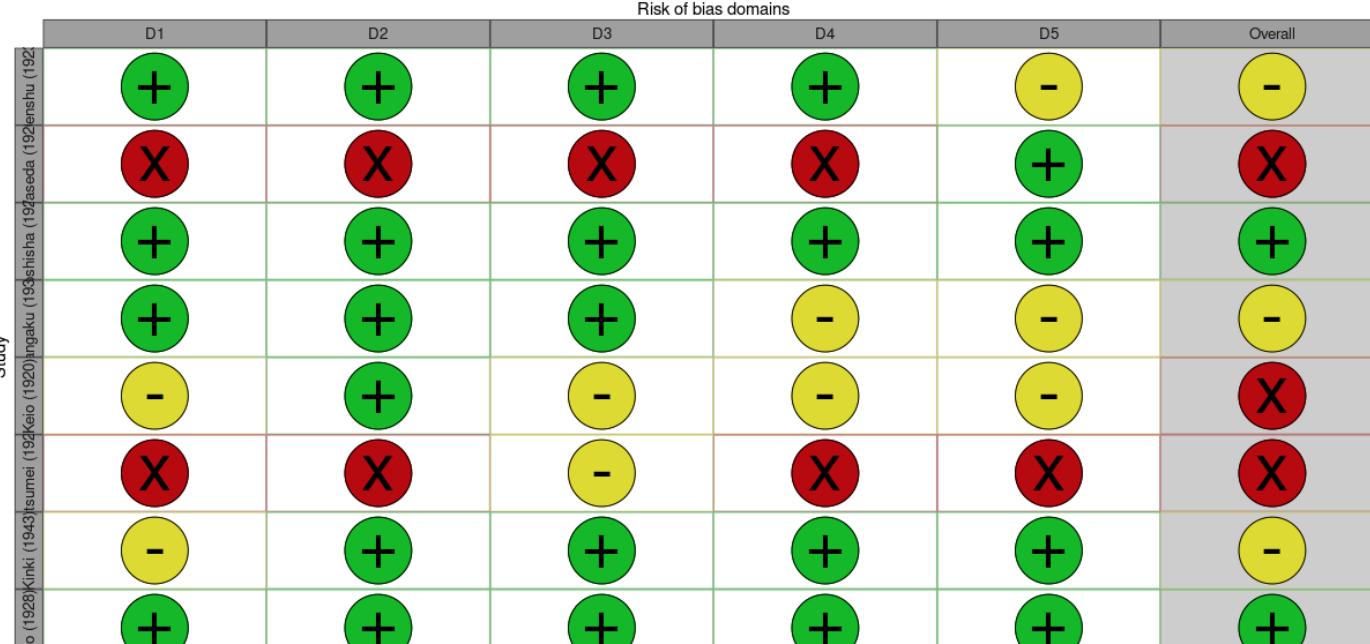
2群の平均値差：RoBプロット(ウェイトなし)

```
rob_summary(data = data_rob, tool = "ROB2", weighted = FALSE)
```



2群の平均値差：個々の研究のRoBプロット

```
rob_traffic_light(data = data_rob, tool = "ROB2")
```



Domains:

- D1: Bias due to randomisation.
- D2: Bias due to deviations from intended intervention.
- D3: Bias due to missing data.
- D4: Bias due to outcome measurement.
- D5: Bias due to selection of reported result.

Judgement

- High
- Some concerns
- Low

2群の平均値差：サブグループ解析

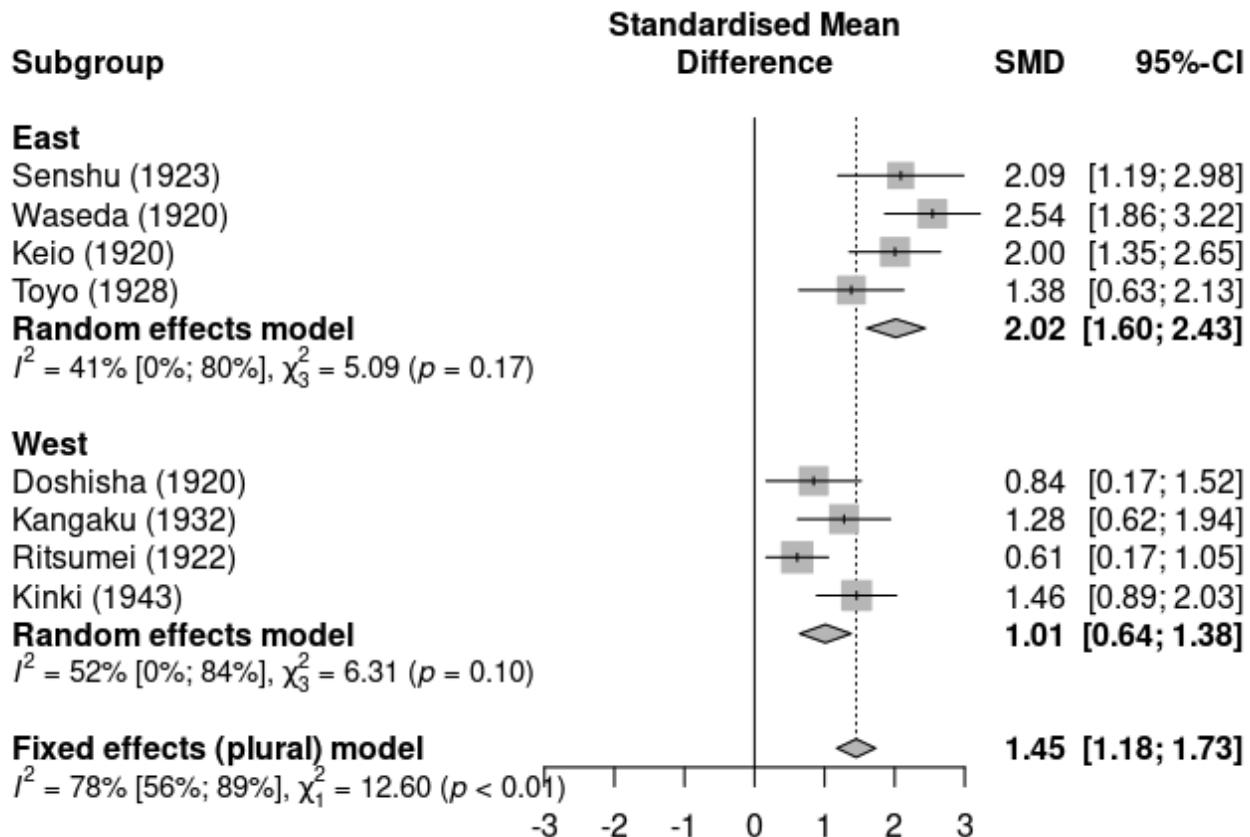
- サブグループ解析は、dmetarパッケージのsubgroup.analysis.mixed.effects関数が便利。

```
result_subgroup <- subgroup.analysis.mixed.effects(result_meta, subgroups = Subgroup)
result_subgroup
```

```
## Subgroup Results:
## -----
##   k   SMD     SE  LLCI  ULCI      p      Q  I2 I2.lower
## East 4 2.017940 0.2127506 1.601 2.435 2.423801e-21 5.085464 0.41      0
## West 4 1.011662 0.1872919 0.645 1.379 6.607663e-08 6.310931 0.52      0
##   I2.upper
## East   0.80
## West   0.84
##
## Test for subgroup differences (mixed/fixed-effects (plural) model):
## -----
##          Q df      p
## Between groups 12.60372 1 0.0003849798
```

2群の平均値差：フォレストプロット

```
forest(result_subgroup)
```



ささやかなTips

- Rでフォレストプロットを作成すると図が途中で切れちゃうことがあります。以下のようにpngで保存する時にwidthとheightを指定するといいです(これは他の図でも同様に使えるTipsです)。

```
png("ファイル名.png", width = 600, height = 600)
forest(メタ分析の結果)
dev.off()
```

その他のRの補足について

以下の内容は、付録に回しました。

- 2値データと相関係数のメタ分析
- マルチレベル分析
- ネットワークメタ分析
- ベイジアンメタ分析
- SEM

系統的展望の手順

1. 研究疑問を作成し、スコープを決める
2. 研究の検索
3. 研究の選択とデータ収集
4. バイアスのリスクの評価
5. メタ分析
6. 結果の評価と解釈

結果の評価と解釈

- 結果の解釈は、 p 値が有意かどうかに焦点をあてない。信頼区間を報告し、エビデンスの確実性について議論をする。
- 統合に用いた効果量だけでなく、解釈しやすい効果量も提示する（2値変数の場合はNNT,連続変数かつ同一尺度の研究の場合はよく使われる単位や最小限の重要な差などを使う）
- GRADEの5つの検討事項（バイアスのリスク、効果の一貫性、不正確性、非直接性、出版バイアス）をもとに、各アウトカムのエビデンス総体の確実性を評価する。エビデンスの確実性について結論をだす。評価には理由を付与する。

まとめ

- ・メタ分析はエビデンスの確実性の評価において有用だが、適切な系統的展望を行うのは容易ではない
- ・系統的展望を行う上で、PRISMAなどの報告ガイドラインやコクランなどの指針が有用である
- ・メタ分析も事前登録が重要であり、事前に研究検索・選択・データ抽出から解析までプロトコルを作ることが推奨される
- ・研究検索・選択・データ抽出の信頼性を高める取り組みもあり、便利なツールを活用する
- ・メタ分析を行う便利なRパッケージが用意されている

謝辞

本発表は、JSPS科研費JP20K20870の助成を受けたものです。

付録：2値変数のメタ分析(データ準備)

2値変数のメタ分析では、介入・実験群のイベント数、介入・実験群のサンプルサイズ、統制群のイベント数、統制群のサンプルサイズ、研究名、（必要なら）研究の特徴の情報（以下では、西か東か）を用意する。以下は適当に作ったデータです。

```
# 使うパッケージ
library(dmetar)
library(meta)

# データを作る (順番に, 相関, 参加者数, 研究, グループ)
Ee <- c(16, 16, 50, 5, 80, 9, 17, 30)
Ne <- c(30, 28, 101, 23, 302, 20, 31, 101)
Ec <- c(15, 14, 45, 10, 120, 11, 15, 40)
Nc <- c(31, 27, 102, 21, 305, 23, 29, 98)
study <- c("Senshu (1923)", "Waseda (1920)", "Doshisha (1920)", "Kangaku (1932)", "Keio (1920)", "R")
group <- c("East", "East", "West", "West", "East", "West", "West", "East")
data_bin <- data.frame(Ee, Ne, Ec, Nc, study, group)
```

付録：2値変数のメタ分析

metaパッケージのmetabin関数を使う。smは"OR"("RR", "RD", "ASD", "DOR"もある)。method.tauも選べるので、今回もML。

```
# メタ分析の実施
result_bin <- metabin(Ee,
                       Ne,
                       Ec,
                       Nc,
                       study,
                       data = data_bin,
                       sm = "OR",
                       comb.fixed = TRUE,
                       comb.random = TRUE,
                       method.tau = "ML")
```

付録：2値変数のメタ分析結果

```
> result_bin
```

| | OR | 95%-CI | %W(fixed) | %W(random) |
|-----------------|--------|------------------|-----------|------------|
| Senshu (1923) | 1.2190 | [0.4461; 3.3314] | 4.0 | 8.0 |
| Waseda (1920) | 1.2381 | [0.4276; 3.5846] | 3.5 | 7.2 |
| Doshisha (1920) | 1.2418 | [0.7149; 2.1571] | 13.1 | 18.9 |
| Kangaku (1932) | 0.3056 | [0.0825; 1.1320] | 4.7 | 5.0 |
| Keio (1920) | 0.5556 | [0.3940; 0.7834] | 50.8 | 29.6 |
| Ritsumei (1922) | 0.8926 | [0.2682; 2.9701] | 3.3 | 5.9 |
| Kinki (1943) | 1.1333 | [0.4106; 3.1282] | 4.1 | 7.8 |
| Toyo (1928) | 0.6127 | [0.3407; 1.1018] | 16.5 | 17.5 |

Number of studies combined: k = 8

| | OR | 95%-CI | z | p-value |
|----------------------|--------|------------------|-------|---------|
| Fixed effect model | 0.7280 | [0.5792; 0.9149] | -2.72 | 0.0065 |
| Random effects model | 0.7829 | [0.5733; 1.0692] | -1.54 | 0.1237 |

Quantifying heterogeneity:

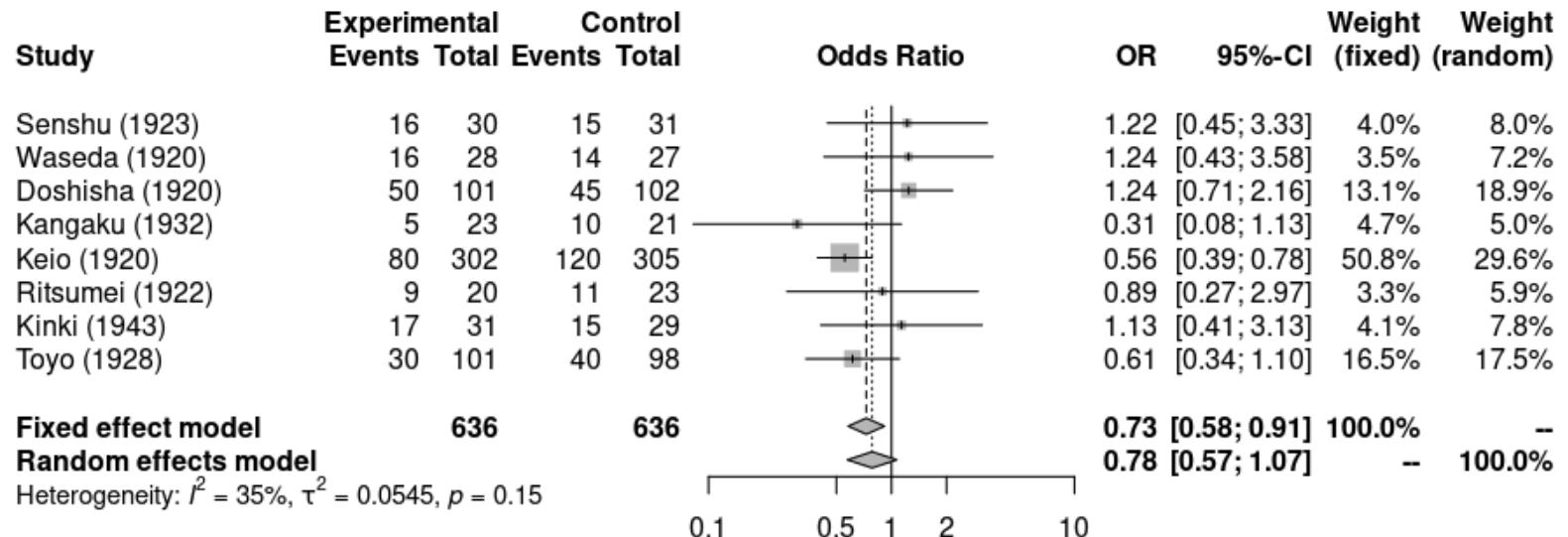
$\tau^2 = 0.0545$ [0.0000; 0.7829]; $\tau = 0.2334$ [0.0000; 0.8848]
 $I^2 = 35.2\%$ [0.0%; 71.3%]; $H = 1.24$ [1.00; 1.87]

Test of heterogeneity:

| Q | d.f. | p-value |
|-------|------|---------|
| 10.80 | 7 | 0.1476 |

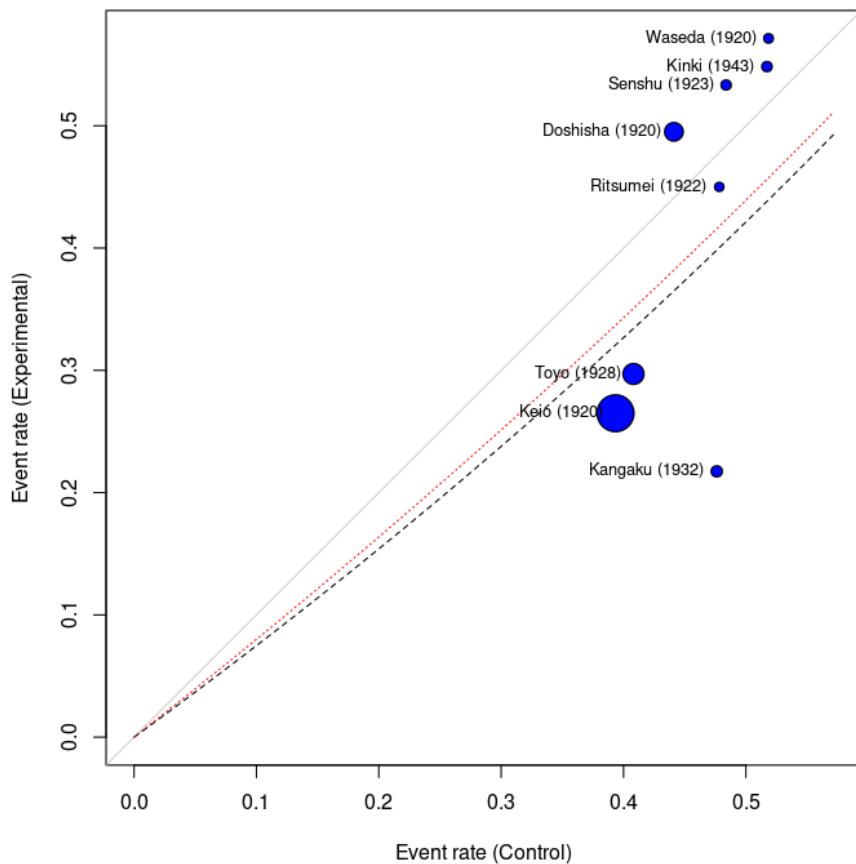
付録：2値変数のメタ分析（フォレストプロット）

```
forest(result_bin)
```



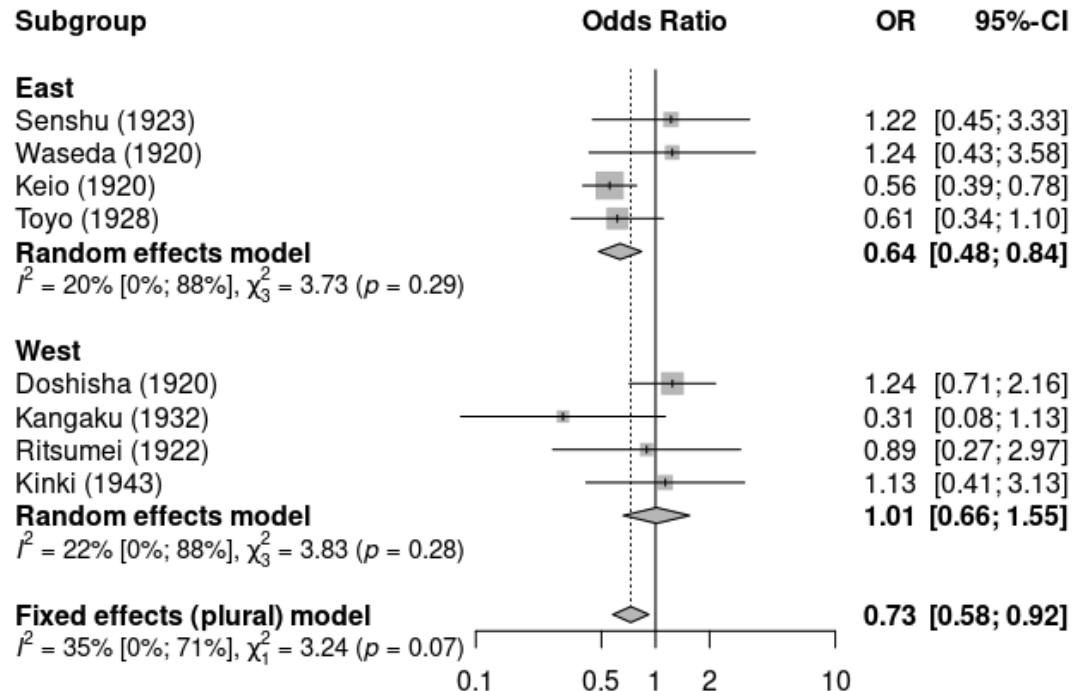
- L'Abbé Plots: 介入・実験群のイベント率と統制群のイベント率をプロットして、サンプルサイズもわかりやすく示したプロット

```
labbe.metabin(result_bin, bg = "blue", studlab = TRUE, col.random = "red")
```



付録：2値変数のメタ分析(サブグループ解析：東西)

```
result_subgroup_bin <- subgroup.analysis.mixed.effects(result_bin, subgroups = group)
forest(result_subgroup_bin)
```



付録:相関係数のメタ分析(データ準備)

相関係数のメタ分析では、各研究の相関係数、サンプルサイズ、研究名、（必要なら）研究の特徴の情報（以下では、西か東か）を用意する。以下は適当に作ったデータです。

```
# 使うパッケージ
library(dmetar)
library(meta)

# データを作る (順番に、相関、参加者数、研究、グループ)
r <- c(.2,.3,.4,.35,.1,.7,.6,.3)
n <- c(30,28,100,20,300,20,30,100)
study <- c("Senshu (1923)","Waseda (1920)","Doshisha (1920)","Kangaku (1932)","Keio (1920)","R
group <- c("East","East","West","West","East","West","West","East")
```

付録:相関係数のメタ分析

metaパッケージのmetacor関数を使う。smは"COR"("ZCOR"もある)。
method.tauも選べるので、今回もML。

```
# メタ分析の実施
result_cor <- metacor(r,
                       n,
                       study,
                       sm = "COR",
                       comb.fixed = TRUE,
                       comb.random = TRUE,
                       method.tau = "ML")
```

付録:相関係数のメタ分析結果

```
> result_cor
```

| | COR | 95%-CI | %W(fixed) | %W(random) |
|-----------------|--------|--------------------|-----------|------------|
| Senshu (1923) | 0.2000 | [-0.1494; 0.5494] | 3.9 | 9.0 |
| Waseda (1920) | 0.3000 | [-0.0432; 0.6432] | 4.1 | 9.2 |
| Doshisha (1920) | 0.4000 | [0.2345; 0.5655] | 17.6 | 15.5 |
| Kangaku (1932) | 0.3500 | [-0.0446; 0.7446] | 3.1 | 7.9 |
| Keio (1920) | 0.1000 | [-0.0122; 0.2122] | 38.3 | 17.5 |
| Ritsumei (1922) | 0.7000 | [0.4707; 0.9293] | 9.2 | 13.0 |
| Kinki (1943) | 0.6000 | [0.3671; 0.8329] | 8.9 | 12.9 |
| Toyo (1928) | 0.3000 | [0.1207; 0.4793] | 15.0 | 15.0 |

Number of studies combined: k = 8

| | COR | 95%-CI | z | p-value |
|----------------------|--------|------------------|------|----------|
| Fixed effect model | 0.3020 | [0.2326; 0.3714] | 8.53 | < 0.0001 |
| Random effects model | 0.3663 | [0.2229; 0.5097] | 5.01 | < 0.0001 |

Quantifying heterogeneity:

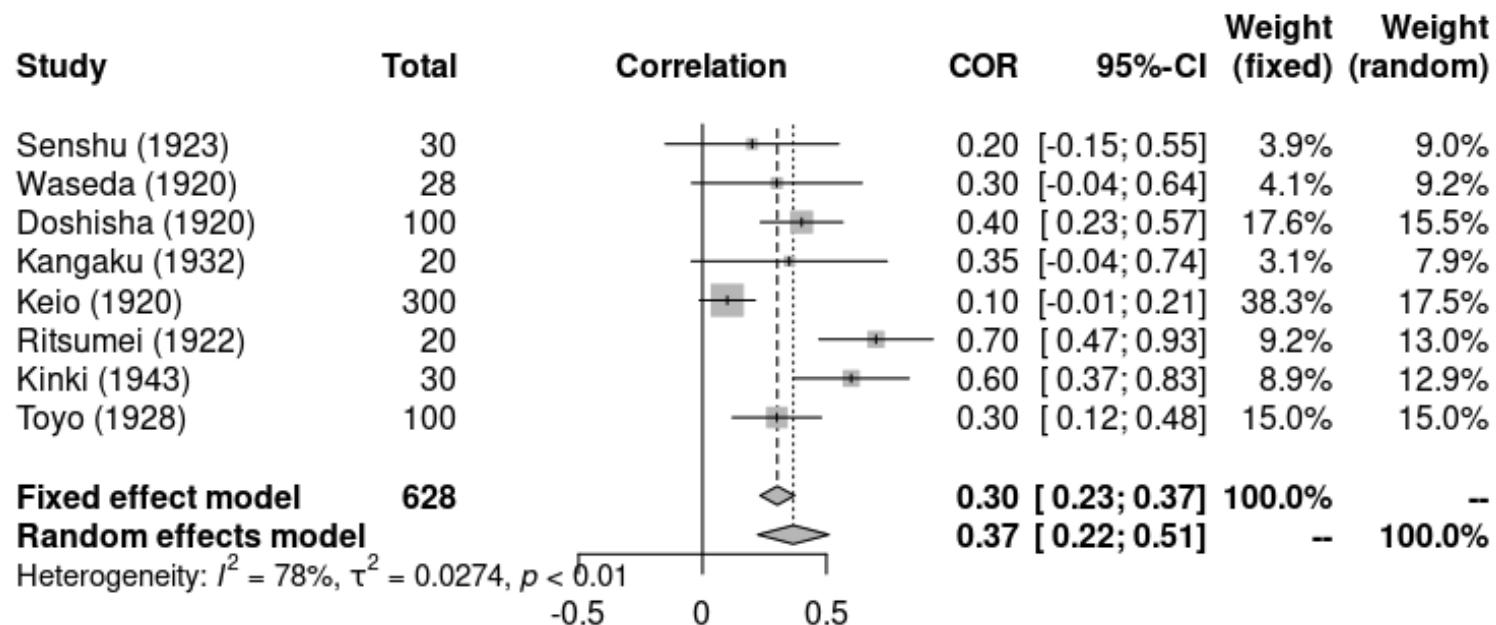
$\tau^2 = 0.0274$ [0.0070; 0.1498]; $\tau = 0.1654$ [0.0834; 0.3870]
 $I^2 = 78.2\%$ [57.0%; 88.9%]; $H = 2.14$ [1.52; 3.00]

Test of heterogeneity:

| Q | d.f. | p-value |
|-------|------|----------|
| 32.04 | 7 | < 0.0001 |

付録:相関係数のメタ分析（フォレストプロット）

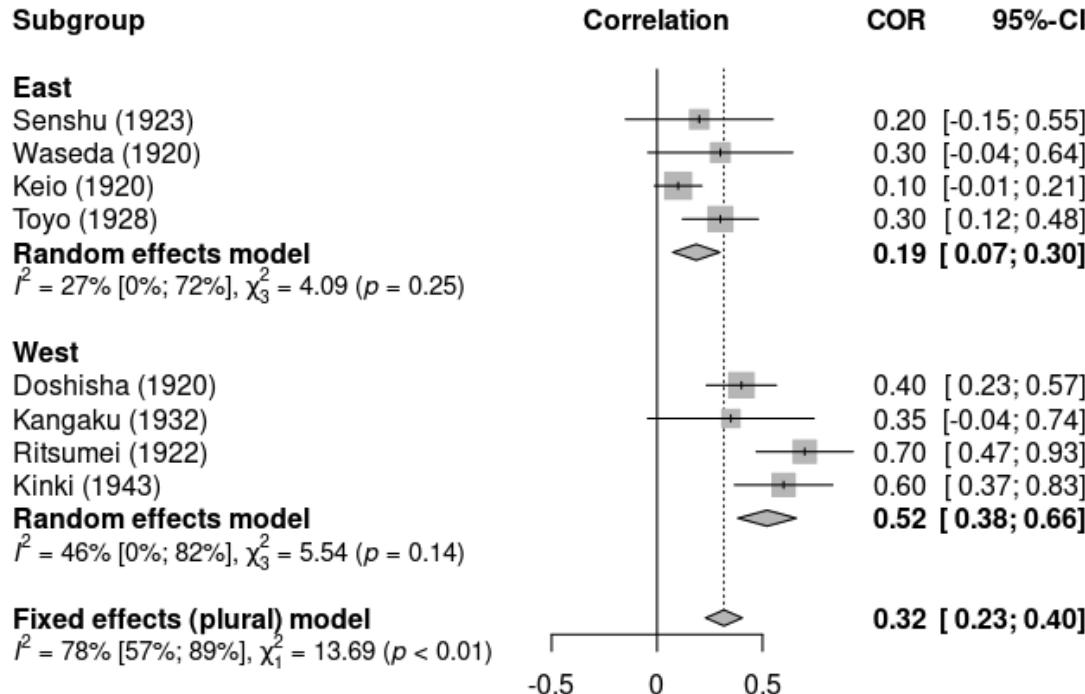
```
forest(result_cor)
```



付録:相関係数のメタ分析(サブグループ解析：東西)

サブグループ解析

```
result_subgroup <- subgroup.analysis.mixed.effects(result_cor, subgroups = group)  
forest(result_subgroup)
```



付録：ベイジアンメタ分析とネットワークメタ分析

- ・ネットワークメタ分析は、2つ以上の介入の効果をネットワークとして検討し、直接比較だけなく、間接比較も検討できる（治療法のランキングも出せる）。
- ・ネットワークメタ分析はnetmetaパッケージで実行できるが、柔軟なモデルを使う場合は、ベイジアンネットワークメタ分析が便利（gemtcパッケージやRStanなどを使う）
- ・以下過去にまとめた資料
 - ベイジアンメタ分析のためのDockerファイル作成法
 - Stanでベイジアンメタ分析（一対比較の固定効果モデル）
 - 入門資料のリンク
 - Stanでネットワークメタ分析（固定効果モデル）

付録：マルチレベルメタ分析

- そもそもメタ分析はマルチレベルな性質がある。ランダム効果モデルには、参加者（レベル1）と研究（レベル2）の2つの階層が含まれている。
- マルチレベルメタ分析という時は、3レベル以上のマルチレベルになる。例えば、参加者が各研究にネストし（レベル1）、いくつかの研究が特定の文化的背景・言語にネストし（レベル2）、それぞれの文化背景が全体の効果の分布にネストする。
- サブグループ分析もマルチレベルで実施でき、ばらつきがどのレベルの変動によって説明できるのかを検討可能
- 分析にはmetaforパッケージが使える。詳しくは、[Doing Meta-Analysis in R](#)参照

付録：MASEM(Meta-analytic structural equation modeling)

- ・メタ分析をSEMの枠組みで実行することができる。
- ・SEMを活用することで、多変量メタ分析ができる。例えば、臨床試験には複数のアウトカムが設定されているが、アウトカム間に関連があることもあり、それを考慮した上で効果を検討することができる。
- ・メタ分析によって複数の研究で報告されている相関行列を統合し、SEMで確認的因子分析を実施したり、媒介分析を実施したりもできる。
- ・詳しくは、Jak(2015), Cheung(2015), 宇佐美(2018)を参照
- ・metaSEMパッケージが使える。

引用文献

- Cheung, M. W. L.(2015). *Meta-Analysis: A Structural Equation Modeling Approach*. Wiley.
- Cooper, H. M.(2020). *Reporting Quantitative Research in Psychology (How to Meet APA Style Journal Article Reporting Standards, Second Edition, Revised)* . American Psychological Association.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine*, 6(7), e1000097.
- Greenhalgh, T. (1997). How to read a paper. Getting your bearings (deciding what the paper is about). *BMJ* , 315(7102), 243–246.
- Higgins, J., Lasserson, T., Chandler, J., Tovey, D., Thomas, J., Flemyng, E., Churchill, R. (2020) *Methodological Expectations of Cochrane Intervention Reviews(MECIR):Standards for the conduct and reporting of new Cochrane Intervention Reviews, reporting of protocols and the planning, conduct and reporting of updates*. サイトへのリンク 日本語翻訳版へのリンク
- Harrer, M., Cuijpers, P., Furukawa, T.A, & Ebert, D. D. (2019). Doing Meta-Analysis in R: A Hands-on Guide. https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/.
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ* , 327(7414), 557–560.
- Higgins, J. P. T., & Thomas, J. (Eds.). (2019). *Cochrane Handbook for Systematic Reviews of Interventions*. Wiley-Blackwell.
- Jak, S. (2015). *Meta-Analytic Structural Equation Modelling*. Springer.
- 国里愛彦 (2015). 系統的展望とメタアナリシスの必須事項 行動療法研究, 41(1), 3-12.
- 宇佐美 慧(2018). 構造方程式モデリングを用いたメタ分析 岡田涼・小野寺孝義（編）『実践的メタ分析入門』 pp.55-66.

引用文献

- ・ 奥村泰之(2014). 非薬物療法の介入研究の必須事項. 行動療法研究, 40(3), 155–165.
- ・ Simonsohn, U., Nelson, L.D., & Simmons, J.P. (2014). P-Curve and Effect Size: Correcting for Publication Bias Using Only Significant Results. Perspectives on Psychological Science, 9, 666 – 681.
- ・ 竹林由武(2014). 観察研究の必須事項. 行動療法研究, 40(3), 167–175.
- ・ 卓興鋼・吉田佳督・大森豊緑 (2011). エビデンスに基づく医療（EBM）の実践ガイドラインシステムティックレビューおよびメタアナリシスのための優先的報告項目（PRISMA声明）. 情報管理, 54(5), 254–266.
- ・ 谷晋二(2015). 症例研究の必須事項. 行動療法研究, 41(1), 13–18.
- ・ 友利幸之介・澤田辰徳・大野勘太・高橋香代子・沖田勇帆 (2020). スコーピングレビューのための報告ガイドライン 日本語版：PRISMA-ScR. 日本臨床作業療法, 7, 70–76.
- ・ 土屋政雄(2015). 尺度研究の必須事項. 行動療法研究, 41(2), 107–116.
- ・ 山田・井上(2012). 『メタ分析入門』 東京大学出版会