

・ システマティックレビュー概論 第4回

2025.1.10

静岡社会健康医学大学院大学

客員教授

公益財団法人 日本医療機能評価機構

客員研究主幹

森實敏夫



# バイアスリスクの評価

1

システマティックレビューでは採用された研究を個別に評価する必要があります。バイアスリスクの評価について解説します。

メインスライドと解説のPDFファイルはこちらURL:

<https://info.zanet.biz/lec/srsz/1-rob-main/1-rob-note.pdf>

このウェブページのURL:

<https://info.zanet.biz/lec/srsz/1-rob-main/one.htm>

RoB2用評価シートはこちら :

<https://info.zanet.biz/lec/srsz/1-rob-main/rob2-sheet.xlsx>

RoB2 Web Tool-Jはこちら :

[https://stat.zanet.biz/sr/crobt2\\_j.htm](https://stat.zanet.biz/sr/crobt2_j.htm)

バイアスリスク評価に対するMindmapはこちらURL: [https://info.zanet.biz/lec/kng/bias\\_risk.htm](https://info.zanet.biz/lec/kng/bias_risk.htm)

(目次にそれぞれのスライドと解説へのリンクが付いています。Quizへのリンクもあります。)

\* ナビゲーションパネルの中央のボタンをクリックすると関連事項、Quizへのリンクが表示されます。

学習ゴールはランダム化比較試験のバイアスリスク評価ができるようになることです。

学習アウトカムは

ランダム化比較試験のバイアスリスク評価ができるようになる。

1. バイアスとバイアスリスクを説明できる。
2. Cochrane Risk of Bias Tool ver.2.0によるバイアスリスク評価ができる

1. バイアスとバイアスリスクを説明できるようになることと、2. Cochrane Risk of Bias Tool ver.2.0によるバイアスリスク評価ができるようになることです。

数多くのバイアスが知られています。バイアスは効果推定値に対して過大評価に働く場合と過小評価に働く場合があります。

なぜバイアスではなくバイアスリスクを評価するのでしょうか？

# バイアスとバイアスリ スク

3

バイアスは「研究結果の系統的な偏り、あるいは、推定の真実からの系統的な偏り」と定義されます。

研究のさまざまな段階でおきる誤差により真の値からずれが生じることが起きます。

バイアスは3つに分類されます：

選択バイアス Selection bias

情報バイアス Information bias

交絡バイアス Confounding bias

研究のデザイン、セッティング、実施の過程、データ分析の過程などさまざまな段階でバイアスが生じる可能性があります。

交絡バイアスはデータ分析の時点である程度は調整が可能です。

# 研究における誤差・バイアス

	研究の段階	誤差			
		偶然誤差	系統誤差 (バイアス)		
			交絡以外のバイアス		交絡
			選択バイアス	情報バイアス	
データそのものの問題	研究対象者の選択	○	○		
	要因・比較対照群の設定	○	○		
	アウトカムの測定	○		○	
データの解釈の問題	効果の推定	○			○

福原俊一 著：臨床研究の道標 第2版、2013年

4

バイアスには数多くの種類があり、特に観察研究の場合は、バイアスの影響を受ける可能性が高くなります。さらに、未知のバイアスの影響、あるいはバイアスの原因となる因子が未測定のこともあります。

どのような研究デザインでもバイアスが生じます。ランダム化比較試験であっても、さまざまなバイアスの影響を受ける可能性があります。また、盲検化することですべてのバイアスが防止できるわけではありません。

# 研究のタイプとバイアスの種類

選択バイアス	情報バイアス	バイアス	メカニズム	研究のタイプ			
				記述研究	コホート研究	症例対照研究	ランダム化比較試験
○		自己選択バイアス	自己判断で研究に参加することが要因曝露とアウトカムの両方に関係がある	○	○		○
○		未回答者バイアス	未回答者より回答者の方が健康状態が良い/その逆	○	○		○
○		ランダムでない脱落	健康状態のよくない人が脱落しやすい/副作用が強い人が脱落する		○		○
○		サンプリングバイアス	源泉集団を代表しない標本集団			○	
○	○	診断バイアス	要因曝露を知っていることでアウトカムの測定（診断）に偏りが生じる		○	○	○
○	○	発見徴候バイアス	要因曝露がアウトカム測定の機会を増やす		○	○	○
	○	ホーンソン効果	研究に参加し、観察されることで態度を変える	○	○		○
	○	面接者バイアス	面接者によって応答が異なる/研究者が対象者から情報を聞き出す際に特定の方向性持つ	○	○	○	○
	○	思い出しバイアス	過去の要因曝露を思い出すことに群間で差がある			○	

福原俊一 著：臨床研究の道標 第2版、2013年より改変

5

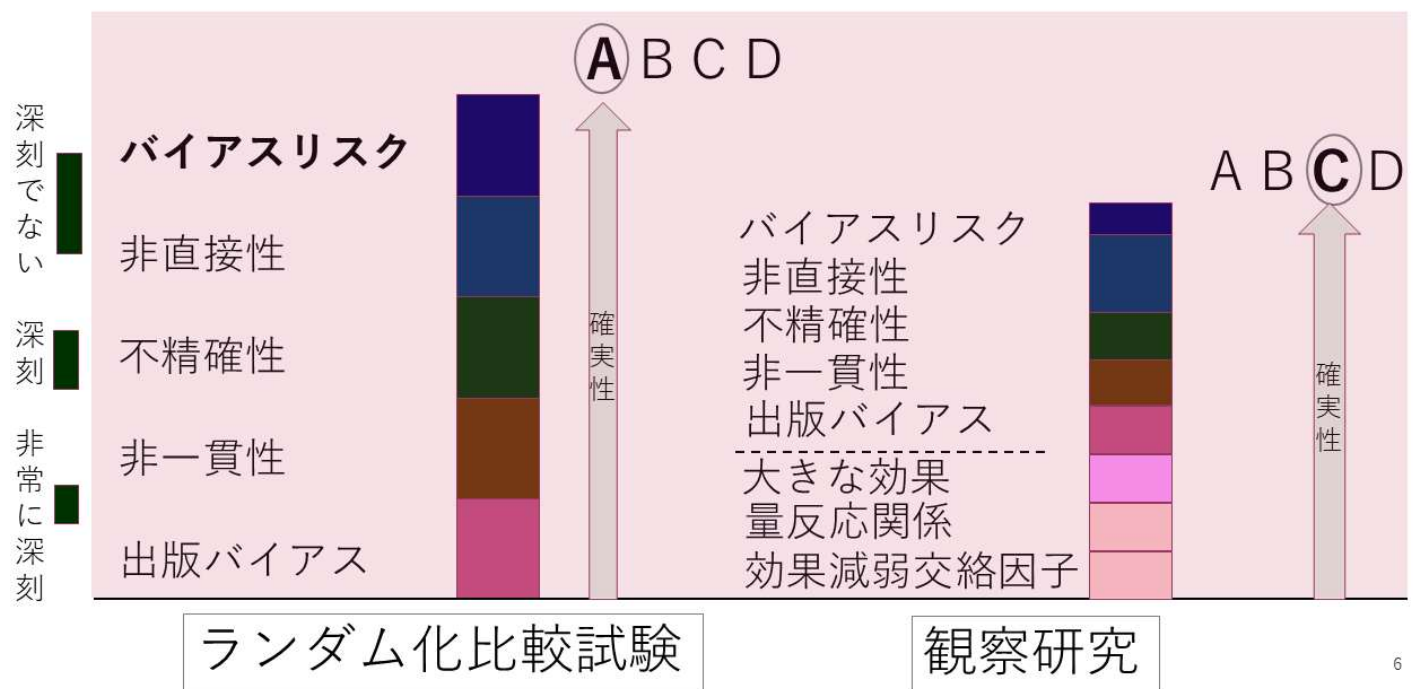
システマティックレビューにおけるエビデンスの評価では、複数の研究をまとめたエビデンス総体のエビデンスの確実性を評価するのが目的になります。エビデンス総体のエビデンスの確実性は、複数のドメインの評価に基づいて行われます。バイアスリスクは、そのうちの一つにあたります。

左側にはランダム化比較試験の場合に評価される5つのドメインすなわちバイアスリスク、非直接性、不精確性、非一貫性、出版バイアスを示します。これらのドメインの評価がそれぞれエビデンスの確実性を支持する程度を合計したものが、エビデンス総体のエビデンスの確実性に相当します。これら5つのドメインに、エビデンスの確実性を低下させるような問題が認められない場合は、エビデンス総体のエビデンスの確実性はA,B,C,Dの4グレードの内、Aと判定されます。

右側には、観察研究の場合の評価ドメインを示します。ランダム化比較試験の場合の5つのドメインに加えて、大きな効果、量反応関係、効果減弱交絡因子の3つのドメインが追加され、合計8つのドメインの評価が行なわれます。そして、観察研究の場合は、未知の交絡



## エビデンス総体のエビデンスの確実性に影響する因子のドメイン



因子による影響を常に想定する必要があるため、エビデンス総体のエビデンスの確実性は、A,B,C,DのうちCからスタートすることになります。このようなドメインとエビデンスの確実性の評価方法はGRADEアプローチに沿ったものです。

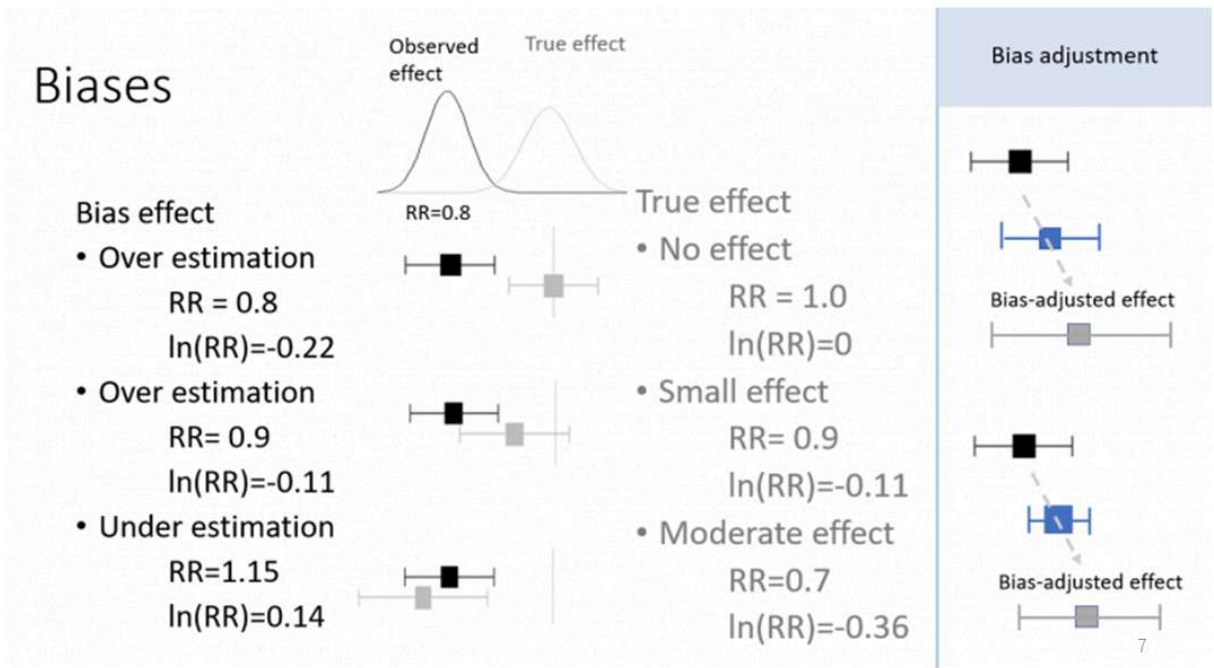
バイアスの効果がどのように研究結果に影響するのかについて考えてみましょう。

その介入には全く効果がないにもかかわらず、あたかも効果があるような結果を得ることがありえます。例えば、有害事象のイベント率が低下するような結果が得られた場合、バイアスが過大評価の効果を示したことになります。この図の中央の上段に示す場合がそれに相当します。グレーで示す真の効果が、バイアスの効果によって、黒で示す結果として現われることになります。

中段に示す例は、真の効果が、小さいにも関わらずバイアスの影響によって、過大評価となり、あたかも大きな効果があるような結果が得られた場合を示します。

下段に示す例は、真の効果は大きいにもかかわらず、バイアスの影響によって過小評価となり、あたかも小さな効果しかないような結果が得られた場合を示します。

# バイアス効果の方向・大きさ・不確実性



多くの場合、バイアスの効果の作用する方向が推定できたとしても、効果の大きさについては正確に決めることは困難です。そのためバイアスのモデルでは、右側に示すように、青で表すバイアスの効果にも点推定値ではなく、一定の分布を想定した不確実性を持たせて考える必要があります。バイアス効果の大きさとその分布を設定することができれば、観察された効果の大きさを調整して真の値を推定することが可能になります。その場合、調整後の効果の大きさの信頼区間はより幅が広くなり確実性が低くなります。

現在主流のバイアス評価法は、バイアスの原因となる事象が起きた可能性をリスクとして評価するだけで、定量的に評価

し、バイアス調整後の真の値を推定することは、ほとんど行われていません。バイアスリスクは効果推定値とは独立して評価されるため、もし、バイアスリスクが深刻であれば、メタアナリシスで得られる統合値と信頼区間は、バイアスの影響を受けた値であって、真の値から偏っているかもしれないということを念頭に解釈する必要があります。

文献 : Turner RM, Spiegelhalter DJ, Smith GC, Thompson SG: Bias modelling in evidence synthesis. J R Stat Soc Ser A Stat Soc 2009;172:21-47. PMID: 19381328

Cochraneのランダム化比較試験に対するRisk of bias toolは2019年に改訂され、バージョン1.1から2.0になりました。

バイアスではなくバイアスリスクという言い方をするのは、先ほど述べたように、バイアスの効果を正確に推定することは多くの場合難しいという考えが背景にあります。一方で、バイアスの効果の大きさとバイアスのリスクつまりバイアスによる偏りが生じている可能性とは別の概念ですから、あらかじめどの程度の大きさのバイアスのリスクを推定するかを決めておく必要があります。Cochrane risk of

# バイアスとバイアスリスク

Cochrane risk of bias tool v.2.0

- 'Risk of bias' is to be interpreted as '**risk of material bias**'. That is, concerns should be expressed only about issues that are likely to affect the ability to draw reliable conclusions from the study.
- Domain-level judgements about risk of bias should have the same implication for each of the six domains with respect to concern about **the impact of bias on the trustworthiness of the result**. A judgement of 'High' risk of bias for any individual domain will lead to the result being at 'High' risk of bias overall, and a judgement of 'Some concerns' for any individual domain will lead to the result being at 'Some concerns', or 'High' risk, overall (see 1.2.3).
- The tool includes **optional judgements of the direction of the bias for each domain and overall**. If review authors do not have a clear rationale for judging the likely direction of the bias, they should not guess it.

Revised Cochrane risk-of-bias tool for randomized trials (RoB 2) (22 August 2019)  
<https://www.riskofbias.info/welcome/rob-2-0-tool/current-version-of-rob-2>

8

bias toolの考え方では、その研究から信頼できる結論を引き出すことが難しいと言えるほどの大きさのバイアスのリスクを評価することになっています。そのような大きさのバイアスのことをmaterial biasという風に呼んでいます。しかしながらどの程度の大きさのバイアスをmaterial biasとするかは明確にされていないことがほとんどです。

また、バイアスのドメインは複数あるわけですが、それぞれのドメインの判定が、同じ程度の大きさのバイアスのリスクを評価するという前提としています。つまりあるドメインのバイアスリスクを高と判定した場合と、別のドメインのバイアスのリスクを高と判定した場合は、いずれも効果推定値は同じ程

度偏っていると考えるということになります。したがって、いずれかのドメインが高の判定になると、その研究のバイアスリスクは高という判定になります。

しかし、ここで問題になるのは、バイアスの効果の方向性、つまり過大評価なのか過小評価なのかを考慮していないということです。ひとつのドメインのバイアスが過大評価に働いて、別のドメインのバイアスが過小評価に働いている場合は、研究で得られた効果推定値は真の値からずれていないことになります。しかし、バイアスの効果の方向性を考慮しないとその研究はバイアスリスクが高いという判定になり確実性が低いとみなされることになってしまいます。Cochrane risk of bias toolはこのバイアスの効果の方向性の評価がオプションになっています。その理由はその評価にはその疾患に対する専門的な知識や経験が必要で、多くのレビュアーにとってその評価が困難なためと考えられます。

わが国では疾患専門家がシステマティックレビューを行うことが多いので、バイアスの効果の方向性については特別な作業をしなくてもある程度評価が可能ではないかと考えられます。しかしながら、あくまでオプションな項目であることから、多くの場合、バイアス



効果の方向は評価されないと思われます。

Revised Cochrane risk-of-bias tool for randomized trials (RoB 2) (22 August 2019)

<https://www.riskofbias.info/welcome/rob-2-0-tool/current-version-of-rob-2>

# 介入研究のバイアスリスク

## ドメイン1 選択バイアス

- ランダム化（ランダム配列の生成、ベースラインの不均衡）
- コンシールメント

## ドメイン2 実行バイアス

- 盲検化（医療者と患者）

## ドメイン3 検出バイアス

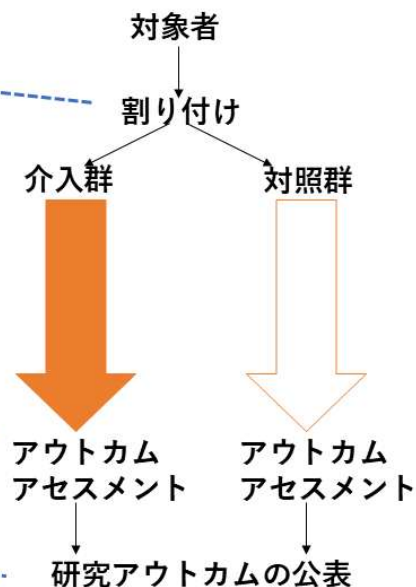
- 盲検化（アウトカム測定者）

## ドメイン4 症例減少バイアス

- （不完全アウトカムデータによるバイアス）
- ITT解析の実施

## ドメイン5 その他のバイアス

- 選択的アウトカム報告
- 早期試験中止
- その他のバイアス



介入研究では介入群と対照群におけるアウトカムが比較され、介入の効果が解析されます。介入群と対照群がどのように設定され、アウトカムがどのように測定されるかによって、研究結果の科学的妥当性は異なってきます。

この図では、介入研究で起きうるバイアスの種類とその対処法、および、どの段階でバイアスが発生するかを示しています。

ランダム化比較試験では、介入群と対照群は介入以外の点で、未知の因子も含めて、アウトカムに影響を与えうる因子はバランスがとれるはずで、もしアウトカムに差が認められれば、それは介入による効果であることを証明することができます。

ランダム化比較試験以外の場合、アウトカムに差が認められても、それが、介入による効果であることを証明することは程度の差はありますが、困難になり、結果には不確実性を伴います。一方、ランダム化比較試験であっても、ランダム割り付けが実現しない場合もあります。例えば、担当医が割り付け表を見ることができると、登録予定の症例の状態に応じて、有効性が高いと思われる治療に登録する操作が可能になり、結果はバイアスの影響を受けたものになってしまいます。これは、コンシールメントが行われなかったためのバイアスになります。

今まで、ランダム化比較試験などの介入研究におけるバイアスとして、選択バイアス、実行バイアス、検出バイアス、症例減少バイアス、その他のバイアス（選択的アウトカム報告など）が評価ドメインとされてきました。そして、選択バイアスはランダム配列の生成、ベースラインの不均衡、コンシールメント、実行バイアスは医療者と患者の盲検化、検出バイアスはアウトカム測定者の盲検化、症例減少バイアスは、不完全アウトカムデータ、Intention-to-treat解析の実施の有無、その他のバイアスは、選択的アウトカム報告、早期試験中止、その他のバイアス、これらについてバイアスリスクとして評価されてきました。

結果に大きな偏りを引き起こすほどの効果をもつバイアスの可能性を、高、中、低の3段階、あるいは高、不明、低の3分類で評価することが一般的でした。

2019年8月にコクランがRisk of bias tool version 2.0へ移行したことに伴い、ランダム化比較試験のバイアスリスク評価法が変わりつつあります。今日は、このRoB2についてウェブツールも用いながら解説します。ただし、ここに示すドメイン1～4のバイアスはRoB2では、名称が変更されていますが、概念は同じままです。

## 個別研究のバイアスリスクの ドメインと項目: RCT

ドメイン	評価項目	Cochrane RoB v.2.0
選択バイアス	①ランダム化： ランダム配列の生成、ベースラインの不均衡 ②割り付けの隠蔽（コンシールメント）	ランダム化の過程から生じるバイス (AHRQ ・ランダム化の過程あるいは交絡で生じるバイス ・研究への参加者の選択におけるバイス)
実行バイアス	③参加者と医療提供者の盲検化	治療企図からの乖離によるバイアス
検出バイアス	④アウトカム測定者の盲検化	アウトカム測定におけるバイアス
症例減少バイアス	⑤ITT解析 ⑥不完全アウトカムデータ	アウトカムデータの欠損のため生じるバイアス
その他のバイアス	⑦選択的アウトカム報告 ⑧早期試験中止 ⑨その他のバイアスの可能性：交差試験のキャリアオーバー、COIによるバイアス、その他	報告結果の選択におけるバイアス

ランダム化比較試験に対する、個別研究のバイアスリスクのドメインと項目について示します。従来、選択バイアス Selection bias、実行バイアス Performance bias、検出バイアス Detection bias、症例減少バイアス Attrition bias、その他のバイアス Other biasesの各ドメインがあり、それぞれのドメインにここに示すような項目が設定されていました。

Cochrane risk of bias tool version 2.0ではこれらのドメインの名称がさまざまな理由から変更されました。しかし、概念はまったく同じです。基本は、ここに示す、“ランダム化の過程から生じるバイアス”、“治療企図からの乖離によるバイアス”、“アウトカム測定におけるバイアス”、“アウトカムデータの欠損のため生じるバイアス”、“報告結果の選択におけるバイアス”の5つに限定されています。

また、各ドメインに対して、いくつかのシグナリングクエスチョンが設定されており、それらにYes, Probably yes, Probably no, No, No informationなどの選択肢から答えを選ぶことで、判定ができるように考案されています。それを実行するExcelマクロも公開されています。ただし、シグナリングクエスチョンは柔軟な取り扱いになっており、あくまでレビューの判定を優先することになっています。

コクラン リスクオブバイアス ツール バージョン 2 ではランダム化比較試験のバイアスを以下の5つのドメインに限定しています。

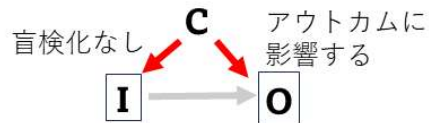
1. ランダム化の過程から生じるバイアス
2. 治療企図からの乖離によるバイアス

# バイアスの分類とRoB 2のバイアスドメイン

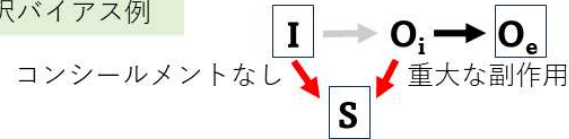
- ① ランダム化の過程から生じるバイアス
- ② 治療企図からの乖離によるバイアス
- ③ アウトカムデータの欠損のため生じるバイアス
- ④ アウトカム測定におけるバイアス
- ⑤ 報告結果の選択によるバイアス

## 交絡バイアス例

- 別の治療を受ける者の割合が多い-②

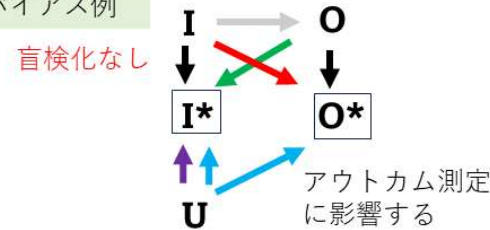


## 選択バイアス例



- 介入群に重症例の割合が多い-①
- 重大な副作用で脱落例が多い-③
- 介入群に多い副作用が報告されない-⑤

## 情報バイアス例



- 望ましいアウトカムの者の割合が多い-④

11

3. アウトカムデータの欠損のため生じるバイアス

4. アウトカム測定におけるバイアス

5. 報告結果の選択によるバイアス

ここでは、バイアスのより基本的な点について解説します。今回のセミナーではオプションな内容ですので、スキップしても構いません。

バイアスの解析を行う際に、Directed Acyclic Graph (DAG) 有効巡回グラフがよく用いられます。DAGでは矢印は因果関係を表し、I, O, Cなどは変数を表していることをまず確認しておきましょう。

例えば、介入を表す変数IはI=0が対照、I=1が介入とします。アウトカムを表す

Oは二値変数としてO=0がアウトカムが生起しなかった、O=1がアウトカムが生起したことを表すとします。そうすると、各群の絶対リスクとリスク比は以下の式で表されます。

$$\text{対照群の絶対リスク } P(O=1|I=0) = P(I=0|O=1) * P(O=1) / P(I=0)$$

$$\text{介入群の絶対リスク } P(O=1|I=1) = P(I=1|O=1) * P(O=1) / P(I=1)$$

$$\text{リスク比 } RR = P(O=1|I=1) / P(O=1|I=0)$$

上記のランダム化比較試験の5つのバイアスがI. 交絡バイアス、II. 選択バイアス、III. 情報バイアスのどれに分類されるか考えてみましょう：

I. 交絡バイアスの例：医師、参加者に対する盲検化が行われない場合。

もし参加者に対する盲検化がないと、参加者は割り付けられた治療が何かを知ることができ、対照群に割り付けられた人は特に、別の治療を受ける割合が多くなったとします。

別の治療を受ける  $C = 1$  ; 別の治療を受けない  $C = 0$  とすると  $P(C=1)$ 、つまり、別の治療を受ける確率が増加します。

その結果は、介入を表す変数  $I$  を変えてしまい、介入の効果が変わってしまうでしょう。

さらに、交絡因子である  $C$  はアウトカム  $O$  にも影響を与えます。実際に受けた別の治療が何かを知ることが多くの場合困難ですが、その治療の効果が  $O$  に対して現れる可能性があります。例えば、もし効果のある既存の治療を受けたとすると、その効果が現れます。

つまり、“別の治療を受ける”という“治療企図からの乖離”は対象者の介入の内容を変え、治療の効果 = アウトカムを変えてしまうので、交絡因子になります。それにより生じるバイアスは交絡バイアスになります。

## II. 選択バイアスの例：コンシールメントがない場合。

コンシールメントがなかった場合、担当医師は次の参加者の割付を知ることができます。介入が有効性が高い新しい治療法の場合、もし重症度の高い患者に対してそれを提供したいと考えたとします。次の割付が対照だった場合、それをパスしてより軽症の患者に参加してもらい、割付を調整することが可能です。そうすると、対象者の選択とくに介入群の選択が影響を受け、介入群は重症度の高い人が多くなります。そのように選択された対象者における介入の効果は真の値からは偏った値になり、過小評価になるでしょう。

また、対象者の選択はアウトカムの影響も受けます。例えば、対照群に比べ、介入群で重大な副作用の頻度が高く、そのために治療を中断する人が多くなると、脱落例が介入群で多くなり、アウトカムデータの欠損のため生じるバイアスが発生します。

このように、対象者の選択は介入とアウトカムの両方の影響を受けた共通結果になります。その場合、 $S$  の値によって分類されるある層だけ、例えば、副作用が軽かった例だけを分析すると選択バイアスが生じます。また、もし選択の変数  $S$  で調整した分析を行っても選択バイアスが生じます。

報告結果を研究者に都合の悪い場合は報告しないようにしたとします。アウトカムが副作用のような有害事象だとします。その場合は、介入群の頻度が対照群より多い場合は、都合の悪い結果になります。つまり、そのアウトカムに関して、介入群の絶対リスクが対照群の絶対リスクより高く、リスク比は 1 より大きくなります。もしリスク比が 1 であれば、介入にも対照にもどちらが望ましいとも望ましくないとも言えません。変数  $S$  は報告するアウトカムの対象例に選択くする場合は  $S = 1$ 、報告しないアウトカムの対象例に選択する場合は  $S = 0$  とします。



リスク比RRはI=1の場合にO=1となる確率をI=0の場合にO=1となる確率を割り算した値になります。

$P(O=1|I=1) = P(I=1|O=1)P(O=1)/P(I=1)$  --- ① この式はベイズの定理です。|は右側の条件が満たされたときに左側が起きる確率、つまり条件付き確率を表します。

$P(O=1|I=0) = P(I=0|O=1)P(O=1)/P(I=0)$  --- ②

$RR = \text{①}/\text{②} = P(O=1|I=1)/P(O=1|I=0) = [P(I=1|O=1)/P(I=0|O=1)] * [P(I=0)/P(I=1)]$

$P(I=0) = 1 - P(I=1)$

そのアウトカムの絶対リスクとRRの計算には、I=0あるいはI=1となる確率の値とO=1となる確率の値が必要ですから、S=1からS=0かを決めるのに変数Iと変数O（図の中のO\*）が影響を与えることがわかります。すなわち、SはIとO\*の共通結果です。したがって、報告結果の選択によるバイアスも選択バイアスに分類できます。

III. 情報バイアスの例について考えてみましょう。

例えば、アウトカム測定者の盲検化が行われていない場合、アウトカム測定者が介入治療に大きな期待を抱いていると、介入治療を受けている参加者の方に望ましい結果になるような測定あるいは誤分類を行う可能性が高まります。その結果、介入群に有利な結果、すなわち過大評価のバイアスが生じるでしょう。いわゆる検出バイアスです。赤い矢印がその関係を示します。

Cochrane risk of bias tool 2.0によるバイアス評価です。

# Cochrane Risk of Bias Tool ver.2.0によるバイアス評価

12

ここから、実際のRoB2の各ドメインとシグナリングクエスチョンについて解説します。

乱数表またはコンピュータ生成乱数を用いて、単純ランダム化（各アームの症例数をコントロールしない）、ブロックランダム化（各アームの症例数をコントロールする）、層別ランダム化、・最小化法（アウトカムに影響する因子でバランスをとる）を用いるのが一般的で、これらの方法であれば、1.1はYesとなります。

コイントス、サイコロ、封筒のシャッフルはYesですが、カルテ番号や誕生日の最後の数字による割り付けであれば、Noとなります。

割り付けを予測できないようにすることがコンシールメントで、例えば、中央管理割り付け、中身を知ることができず開封を知ることができる封筒法などであれば1.2はYesとなります。治療薬は連続数が付けられ、中身を推測できない形で管理される必要があります。

# ドメイン1 ランダム化の過程から生じるバイアス スリスク (選択バイアス)

”研究対象の選択の偏りにより生じるバイアス。”

1.1 割り付け配列はランダムか? <input type="text" value="Yes"/>
1.2 割り付け配列は参加者がエンロールし介入に割り付けられるまで隠蔽(concealment)されていたか? <input type="text" value="Yes"/>
1.3 介入群間のベースラインの不均衡はランダム化の過程の問題を示唆するか? <input type="text" value="No"/>

ランダム配列  
の生成

コンシールメ  
ント

ベースライン  
の不均衡

Patients' profileの表で、ベースラインの不均衡について、検討します。比較される群間で、異なる属性がある場合、有意差が無くても、アウトカムに影響する場合があります。その因子は交絡因子として結果に影響を与える可能性があります。ランダム化に問題があったことを示唆します。逆に、属性に過剰な一致がある場合は、群間で人為的な調整が行われたことが疑われます。症例数に応じたランダムな、ずれがあるのが正常です。

ベースラインの不均衡があるため、多変量解析などで調整された場合、ランダム化に問題があるとみなすことが望ましいとされています。アウトカムに影響を与える、すべての因子が測定されていて、それらで調整することはほぼ不可能

13

なので、その研究は非ランダム化比較試験とみなし、ROBINS-Iなどで評価すべきとされています。

Web tools for systematic reviews and for developing clinical practice guidelines

<https://stat.zanet.biz/sr/>

RoB2 Web Tool-J

[https://stat.zanet.biz/sr/crobt2\\_j.htm](https://stat.zanet.biz/sr/crobt2_j.htm)

2.1、2.2のシグナリングクエスチョンは参加者の盲検化、医療者の盲検化に関するものです。これらがYesであるだけでは、判定は確定しません。

## ドメイン2 治療企図からの乖離によるバイアス リスク（実行バイアス）-1

“介入の実行方法の偏りによって生じるバイアス。”

●Part 1:

2.1 参加者は試験期間中割り付けられた介入に気付いたか？

2.2 介入を提供する医療者は試験期間中参加者に割り付けられた介入に気付いたか？

2.3 2.1または2.2でY/PY/NIの場合:企図された介入からの乖離は試験の文脈から生じたか？

2.4 2.3でY/PYの場合:乖離はアウトカムに影響したと思われるか？

2.5 2.4でY/PY/NIの場合:企図された介入からの乖離は群間でバランスがとれているか？

盲検化

介入・ケアの  
実行の群間の  
違い

2.3での“試験の文脈”から生じる乖離とは、“募集と契約による参加者への効果および医療提供者による介入の提供に対する無意識あるいは意識的プロセスへの効果によって生じるもの”をいいます。つまり、参加者は臨床試験に参加することを決め、契約する際に、自分が割り付けられる介入の候補がどのようなものかを知ることになります。そして、自分の受けている介入が何かを知ることができたり、推定できるような場合、他の治療を探し求めて、それを追加で受けたり、割り付けられた介入がプロトコール通りに受けていない事態が発生する可能性が出てきます。治療へのアドヒアランスが変わってくる場合もあります。医療提供者も、割り付けられた介入を知ることができたり、その参加者の治療効果から推

定したりして、他の治療を追加したりすることが起きえます。これらは“試験の文脈”から生じる乖離に相当します。試験の文脈とは関係なく、例えば、何らかのサプリメントを摂取したり、何らかの健康増進の活動をするようなことは含まれません。

2.4ではそのような乖離がアウトカムに影響したと思われるかどうかを問うています。

そして、2.5では、結果として乖離が群間でバランスがとれていると思われるかを問うています。

副作用が出た場合の、治療法の変更がプロトコールで認められている範囲で行われている場合は、治療企図からの乖離とはみなしません。

Part 2では、解析方法について問うています。

## ドメイン2 治療企図からの乖離によるバイアスリスク (実行バイアス) -2

“介入の実行方法の偏りによって生じるバイアス。”

● Part 2:

2.6 介入への割り付けの効果の推定に適切な解析が用いられたか? Yes

2.7 2.6でN/PN/NIの場合:ランダム化された群の参加者の解析の失敗は(結果)に大きな影響を与えるか? NA

介入・ケアの  
実行の群間の  
違いの調整

15

解析の手法として、Intention-to-treat解析、Modified intention-to-treat解析、As-treatedまたはPer-protocol解析などがあります。

Intention-to-treat解析（治療企図解析）、Modified intention-to-treat解析であれば、適切な解析とみなし、Yesとなります。（不完全アウトカムデータへの対処法としては、多変量補完法が望ましいとされています。）

As-treatedまたはPer-protocol解析の場合は、適切な解析とはみなしません。

ランダム割り付け後の適格症例の除外は不適切とみなします。ランダム割り付け後の不適格例の除外は一般的に適切とみなします。

2.7では、除外されたり、企図された介入から乖離した症例がどの程度の割合ならば結果に大きな影響を与えるかについて考えることが求められますが、一様な基準を設定することは難しいとされています。5%程度の割合でもYesとなる場合もあり得ます。

3.1について、連続変数のアウトカムに対しては、参加者の95%のデータが得られれば十分と言えます。二値変数アウトカムの場合は、脱落例数は直接、効果推定値に影響しますが、イベント数が脱落例数よりずっと多い場合は、バイアスは小さいでしょう。

NIはアウトカムデータの欠損について論文に記載がない場合です。この場合は、多くの場合、高バイアスリスクになります。

3.2について、アウトカムデータの欠損によるバイアスを調整する方法が用いられた場合、データ欠損の程度に関する感度分析で結果に大きな変動がない場合はYesまたはProbably yesになります。多変量補完法でアウトカムデータの欠損を補完した場合、介入群だけに適用することは適切ではありません。



## ドメイン3 アウトカムデータの欠損のため生じるバイアスリスク (症例減少バイアス)

解析対象となる症例の減少に系統的な差がある場合に生じるバイアス。

3.1 このアウトカムデータはランダム化されたすべてのあるいはほぼすべての参加者に対して得られたか?

3.2 3.1でN/PN/NIの場合:結果がアウトカムデータの欠損でバイアスを受けていないエビデンスがあるか?

3.3 3.2でN/PNの場合:アウトカムの欠損がその真の値に依存する可能性があるか?

3.4 3.3でY/PY/NIの場合:アウトカムの欠損がその真の値に依存していたか?

### 不完全アウトカムデータによるバイアス

3.3について、アウトカムデータの欠損が参加者の健康状態に関連して起きている場合は、YesまたはProbably yesになります。脱落がアウトカムと関係ない理由によることが記録されている場合は、NoまたはProbably noになり、バイアスリスクは低となります。生存分析での打ち切り例は、解析には含めますが、アウトカムデータの欠損した例とみなしません。

3.4について、3.3がYesの場合、中バイアスリスク (Some concerns) の判定になりますが、3.4がYesの場合は、高バイアスリスクの判定になります。

介入群と対照群の脱落例数に差があり、介入と対照のアウトカムに対する効果に差があり、アウトカムデータの欠損が真

の値に依存する場合は、アウトカムデータの欠損に差がある場合は、バイアスリスクが高まり、Yesとなります。生存分析の場合は、打ち切りの群間差がある場合に同様です。

すでに、同様のアウトカムデータの欠損が真の値に依存するというエビデンスが報告されている場合も、Yesとなります。また、群間でアウトカムデータの欠損が異なる理由が報告されている場合も、Yesとなります。

例えば、統合失調症で報告されているように、症状の持続は、脱落につながるものが広く認められているように、アウトカムデータの欠損が真の値に依存する状況が知られている場合も、Yesです。

生存分析の場合は、参加者が割り付けられた治療を中止あるいは変更した時、例えば、副作用で中止したり、セカンドラインの治療に変更したような場合、打ち切りとなります。

アウトカムデータの欠損と真の値の関係を説明できる参加者の属性で解析を調整できるような場合は、Noです。

## ドメイン4 アウトカム測定におけるバイアスリスク (検出バイアス)

17

### アウトカム測定方法の偏りにより生じるバイアス。

4.1 アウトカム測定法は不適切だったか?

4.2 アウトカム測定・確認は介入群間で異なるか?

4.3 4.1と4.2でN/PN/NIの場合:アウトカム測定者は研究参加者が受けている介入に気づいたか?

4.4 4.3でY/PY/NIの場合:アウトカムの評価は受けている介入を知ることによって影響を受けたことが(今まで)あったか?

4.5 4.4でY/PY/NIの場合:アウトカムの評価は受けている介入を知ることによって影響を受けた可能性があるか?

### アウトカム測定者の盲検化

4.1は、アウトカムの選択ではなく、アウトカム測定法に関するものです。通常、事前に測定が計画されている場合、Noとなります。

測定法の感度の面で介入の効果の範囲外の値を測定しようとする場合、測定機器の妥当性が低い場合はYesまたはProbably yesとなります。

4.2は、比較する群間で、測定方法や、データ収集(タイミングなど)に差があれば、YesまたはProbably yesになります。

4.3は、盲検化されていれば、Noです。なお、患者報告アウトカムの場合は、アウトカム測定者は参加者自身になります。

4.4は、たとえば、全死亡は、影響を受けるアウトカムでないことは知られていますが、患者報告アウトカム(痛みのレベルのような)、ある程度の判断が必要な観察者報告アウトカム、介入提供者の決断アウトカムの場合は、Yesになります。

4.5は、受けている介入を知ることによって、アウトカムの評価が影響を受けたことが分かっているが、それが起きたことを信ずるに足る理由がない場合(Some concernsの判定となる)とアウトカムの評価が影響を受けたと思われる場合(Highの判定になる)があります。介入に有益あるいは有害な効果があることに強い確信があるときは、受けている介入を知ることによって、アウトカムが影響を受ける可能性が高まります。

## ドメイン5 報告結果の選択におけるバイアスリスク (その他のバイアスの内の選択的アウトカム報告)

都合の悪い結果が除外されて測定されたアウトカムの全部が報告されていないため生じるバイアス

5.1 この結果をもたらしたデータは盲検解除前に解析のためにアウトカムデータが得られる前に最終化された事前の解析計画に従って解析されたか?

5.2 評価される数値結果は、そのアウトカムドメインで多数の適格なアウトカム測定（尺度、定義、時点）から、結果に基づいて、選択された可能性があるか?

5.3 評価される数値結果は、結果に基づいて、多数のデータ解析から選択された可能性があるか?

### プロトコール (計画書) と のズレ

5.1解析方法は事前に定められている必要があり、盲検解除後に得られたデータに合わせて変更してはなりません。もし必要な場合でも、解析方法の変更は、盲検解除の前に行うべきです。

5.2は、プロトコールに従ったアウトカムのデータが報告されていれば、NoまたはProbably noとなります。分析の意図が不明あるいは報告されていない場合で、アウトカムドメインの測定法が複数ある場合はNIとなります。

従来、その他のバイスには、以下の項目も含まれていましたが、現在RoB2の基本形に、これらは含まれていません。

○項目：早期試験中止バイアス

計画された分析結果報告か？都合のいい結果が出た時点で中止していないか？

→計画書で予定された分析か？

○項目：その他のバイアス

COIや研究の資金源が結果に偏りを与えないか？


その他に結果を歪めるバイアスは存在しないか？

バイアス効果の方向に対する判定の略語については以下の通りです：

NA Not Applied 適用しない。

# RoB2 Web Tool-J

[https://stat.zanet.biz/sr/crobt2\\_j.htm](https://stat.zanet.biz/sr/crobt2_j.htm)

Study ID:  Outcome:  Menu 

- ランダム化の過程から生じるバイスリスク  NA  fE  fC  tN  aN  uP
- 治療企図からの乖離によるバイアスリスク  NA  fE  fC  tN  aN  uP
- アウトカムデータの欠損のため生じるバイアスリスク  NA  fE  fC  tN  aN  uP
- アウトカム測定におけるバイアスリスク  NA  fE  fC  tN  aN  uP
- 報告結果の選択におけるバイアスリスク  NA  fE  fC  tN  aN  uP

コメント

19

fE Favors the experimental intervention 介入が優位。

fC Favors the comparator 対照が優位。

tN Towards the null effect 無効果の方向へのバイアス。（例：年齢層が幅広いなど、異質性の大きな集団を対象にしている。）

aN Away from the null effect 無効果から離れる方向へのバイアス。（例：均質な集団に偏っている。）

uP Unpredictable 予測不可能。

RoB 2ウェブツール URL:

[https://stat.zanet.biz/sr/crobt2\\_j.htm](https://stat.zanet.biz/sr/crobt2_j.htm)

RoB Web toolの結果をMinds評価シートに転記し、研究ごとのバイアスリスクのまとめの評価を確認します。

もし評価者の評価結果とシグナリングクエスチョンに基づく評価結果が異なる場合は、前者を優先します。そして、その理由を記載しておく必要があります。



# 評価シート

Study ID: Salminen P 2015 Outcome: 1年以内の再発 Menu

-1 -1 0 -2 0 → Copy to Clipboard -1 -1 0 -2 0 -2 → Copy to Clipboard with Summary

- ランダム化の過程から生じるバイアスリスク  NA  FE  FC  TN  aN  uP
- 治療企図からの乖離によるバイアスリスク  NA  FE  FC  TN  aN  uP
- アウトカムデータの欠損のため生じるバイアスリスク  NA  FE  FC  TN  aN  uP
- アウトカム測定におけるバイアスリスク  NA  FE  FC  TN  aN  uP
- 報告結果の選択におけるバイアスリスク  NA  FE  FC  TN  aN  uP

A		B	C	D	E	F	G	H	I	J	K	L	M	N	O							
【SR-5 RoB2 評価シート 介入研究】 : 二値変数 R script exdata																						
診療ガイドライン																						
対象	成人の急性虫垂炎 *各項目の評価は「高(-2)」、「中/疑い(-1)」、まともは「高(-2)」、「中(-1)」、「低(0)」アウトカムごとに別紙にまとめる。																					
介入	抗菌薬投与による保存的治療																					
対照	外科的虫垂切除																					
アウトカム																						
1ヶ月以降1年以内の無再発																						
個別研究																						
研究コード	デザイン	バイアスリスク					非重複性					対照群分子	対照群分母	割合 (%)	分子	分母 (%)	標準差 (種類)	標準差 (値)	信頼区間			
		ランダム化の過程	治療企図からの乖離	アウトカムデータの欠損	アウトカム測定	報告結果の選択	まとも	対象	介入	対照	アウトカム									まとも		
12	Salminen P 2015	RCT	-1	-1	0	-2	0	-2	0	-1	0	0	0	273	272	100%	257	186	72% RR	0.726	0.673 - 0.784	
13	Vons C 2011	RCT	-1	-1	-2	-2	0	-2	0	0	0	0	0	119	104	87%	120	72	60% RR	0.687	0.584 - 0.807	
14	Styrud J 2006	RCT	-1	-1	-1	0	0	-2	-1	-1	-1	0	-1	124	124	100%	128	97	76% RR	0.759	0.688 - 0.838	
15	Eriksson S 1995	RCT	-1	-1	0	-1	0	-2	-1	-1	-1	0	-1	20	20	100%	20	13	65% RR	0.659	0.475 - 0.912	
16																					0.73	0.690 - 0.771
コメント (該当するセルに記入)																						

1.3 割り付けが計画された症例数と実際の登録症例数が異なり、二群でアンバランス。  
 2.1: 2.2: 2.3: 2.5 介入は外科手術なので盲検化はできない。再発の判定の際に影響したかもしれない。  
 4.3 手術を受けたことは知ることができる。  
 4.4 手術を受けていれば再発は除外し、抗菌薬投与による保存的治療群であれば再発をより強く疑う可能性がある。

それぞれの研究について各ドメインの評価をした後、その研究について全体的なバイアスリスクの評価をすることになります。その際のCochrane risk of bias toolの基準をここに示します。

全てのドメインの評価が低の場合には、その研究のバイアスリスクも低、少なくとも一つのドメインの評価が中の場合にはその研究のバイアスリスクは中 (Some concerns)、少なくとも一つのドメインの評価が高の場合にはその研究のバイアスリスクは高、いくつかの複数のドメインの評価が中の場合にはその研究のバイアスリスクは高、というのがおおよその基準になっています。(ここではSome concernsと中は同じ意味で用いています)。ただし、中の評価のドメインがいくつかあったら高とみなすかということについては、それぞれの評価者によります。

アメリカのAgency for Healthcare Research and Quality (AHRQ)は、バイアスの効果推定値の確実性への影響だけでなく、効果推定値に対するバイアスの効果の大きさと方向に対する評価を求めています。



# Overall judgement

Table 1. Reaching an overall risk-of-bias judgement for a specific outcome.

Overall risk-of-bias judgement	Criteria
Low risk of bias	The study is judged to be at <b>low risk of bias for all domains</b> for this result.
Some concerns	The study is judged to raise <b>some concerns</b> in at least one domain for this result, but not to be at high risk of bias for any domain.
High risk of bias	The study is judged to be at <b>high risk of bias</b> in at least one domain for this result. Or The study is judged to have <b>some concerns for multiple domains</b> in a way that substantially lowers <u>confidence in the result</u> .

Cochrane, GRADEのバイアスリスクの評価では、**通常**、バイアスの**効果推定値の確実性**への影響を評価するだけで、効果推定値に対する効果の大きさや方向に対する影響は評価しない。

AHRQは、バイアスの効果推定値の確実性への影響だけでなく、効果推定値に対するバイアスの効果の大きさや方向に対する評価を求めている。

すでにお話ししたように、バイアスの効果の大きさと方向を評価するかしないかによって、それぞれの研究のバイアスリスクの評価が変わってくる可能性があります。バイアスリスクの評価をする場合および評価結果を見る場合には、その点の考慮が必要になります。しかしながら、ほとんどのシステマティックレビューではバイアスの効果の大きさと方向性の評価が行われておらず、バイアスのリスクのみの評価結果が提示されています。

バイアスリスクの各ドメインの評価結果を研究ごとにまとめ、さらに研究横断的に複数の研究をまとめる場合、すなわちエビデンス総体のバイアスリスクは、各研究のサンプルサイズとイベント率を考慮した重みづけを用いて、研究ごとのまとめの評価が大部分の研究で低であればまとめも低（深刻でない）、そして研究ごとのまとめの評価が大部分の研究で中であればまとめも中（深刻な限界）、研究ごとのまとめの評価が大部分の研究で高であればまとめも高（非常に深刻な限界）と判定します。

# バイアスリスク Risk of bias

各研究のサンプルサイズとイベント率を考慮した重み付けを用いて、研究ごとのまとめの評価が大部分の研究で

- 低(0):Lowなら、  
低(0): No serious limitations深刻でない
- 中(-1): Moderateなら、  
中 (-1): Serious limitations深刻な限界
- 高(-2): Highなら、  
高(-2): Very serious limitations非常に深刻な限界と判定する。

研究コード	リスク*					非直接性*		
	削減バイアス	その他	早期中止	その他のバイアス	対照	介入	アウトカム	
Vons C 2011	0	-2	0	0	-1	-1	0	0
Hansson J 2009	0	-2	0	0	-2	-2	-1	0
Styrud J 2006	0	0	0	0	-1	-2	-1	0
Eriksson S 1995	0	0	0	0	-1	-2	-1	0

バイアスリスク

GRADE Table 3 Column 3 JCE ser.4, p411

22

さて、非直接性はクリニカルクエスチョンのPICOと研究のPICOのずれによって生じる不確実性を表すドメインです。

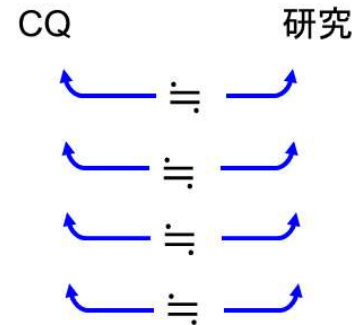
非直接性があると、効果推定値に対して、真の値から偏りを生じ、非直接性はバイアスと同じ効果を持ちます。したがって、外的バイアスExternal biasと呼ぶ研究者もいます。また、適用可能性Applicability、外的妥当性External validity、一般化可能性Generalizabilityと同じ意味で用いられますが、適用可能性は個別患者で判断が必要になるので、AHRQは非直接性を適用可能性と分離する考えを採用しています。

わが国で作成する診療ガイドラインの対象は日本人なので、研究が欧米で行われている研究の場合、人種差が問題になる疾患があります。また、投薬の用量が異なることもあります。同じ疾患でも医療事情の差で、病期や重症度が異なる場合もあります。これらは非直接性として評価されます。

# 非直接性(Indirectness)

クリニカルクエスチョンと研究の間のずれ  
⇒ 本来知りたい効果から偏りを生じる  
=バイアスと同じ効果

- 対象者 P
- 介入 I
- 対照 C
- アウトカム O



\*適用可能性、外的妥当性、一般化可能性、外的バイアスと同じ意味だが、CQの方から見た概念。非直接性がある場合がほとんど。

23

以上述べてきた、バイアスリスクと非直接性は個別の研究について評価を行う必要があります。ここからは、複数の研究を前提にしたエビデンス総体の評価に進むこととなります。

## まとめです

バイアスは研究結果が真の値から偏りを生じさせ、結果の信頼性を低下させる要素で、システマティックレビューではその評価が必須です。

- その研究から信頼できる結論を引き出すことが難しいと言えるほどの、大きさのバイアスのリスクの程度を評価します。
- Cochrane Risk of Bias Tool ver.2.0ではシグナリングクエスチョンに答えていくことでランダム化比較試験のバイアスリスクの評価を行います。
- 結果は各ドメインをLow, Some concerns, Highの3段階で評価し、研究ごとにまとめ、さらに複数の研究の評価をまとめて、エビデンス総体のバイアスリスクの評価とします。

目次パネルの右のボタンをクリックするとQuizへのリンクが表示されます：

## まとめ

- その研究から信頼できる結論を引き出すことが難しいと言えるほどの、大きさのバイアスのリスクの程度を評価する。
- Cochrane Risk of Bias Tool ver.2.0ではシグナリングクエスチョンに答えていくことでランダム化比較試験のバイアスリスクの評価を行う。
- 結果は各ドメインをLow, Some concerns, Highの3段階で評価し、研究ごとにまとめ、さらに複数の研究の評価をまとめて、エビデンス総体のバイアスリスクの評価とする。