

バイアスを可視化する

メタデータ	言語: jpn 出版者: 公開日: 2018-08-17 キーワード (Ja): キーワード (En): 作成者: 森實, 敏夫, 小島原, 典子 メールアドレス: 所属:
URL	http://hdl.handle.net/10470/00032029

バイアスを可視化する

¹公益社団法人日本医療機能評価機構²東邦大学医学部³大船中央病院消化器・IBDセンター⁴東京女子医科大学医学部衛生学公衆衛生学（二）講座モリザネ トシオ コジマハラノリコ
森實 敏夫^{1~3}・小島原典子^{1,4}

（受理 平成29年12月13日）

Visualizing Bias Effects

Toshio MORIZANE^{1~3} and Noriko KOJIMAHARA^{1,4}¹Japan Council for Quality Health Care²Toho University School of Medicine³Ofuna Chuo Hospital⁴Department of Public Health, School of Medicine, Tokyo Women's Medical University

In systematic reviews, the evaluation of the risk of bias and indirectness on effect estimates is essential to assess the certainty of the body of evidence. A program in the programming language R was developed to execute a meta-analysis adjusting for the effects of biases after assessing bias magnitudes, directions, and uncertainty. The program visualizes the effects of biases by creating an overlaid forest plot of bias-adjusted and -unadjusted values. When performing meta-analyses and sensitivity analyses, this helps assess the certainty of the body of evidence by setting various estimates for biases and indirectness of individual studies as risk ratios and standard deviations or 95 % confidence intervals.

Key Words: bias, indirectness, meta-analysis, randomized controlled trial, forest plot

緒 言

診療ガイドライン作成プロセスにおいて研究結果の確実性の評価は重要な要素である¹⁾²⁾。GRADE approach³⁾⁴⁾ではシステマティックレビューにおいてクリニカルクエスチョンに関連して包括的に収集された研究集合について、ランダム化比較試験(RCT)の場合であれば、アウトカムごとにバイアスリスク、非直接性、非一貫性、不精確、出版バイアスの5つのドメインを評価し、エビデンス総体としての確実

性が評価される。観察研究の場合には、効果の大きさ、量反応関係、効果減弱交絡因子について評価される。

臨床研究におけるバイアスについては、Berkman NDらのAHRQの報告書⁵⁾に述べられているように多くの種類がある一方で、経験的なデータに基づいたメタ疫学的研究は少なく、疾患、介入、アウトカムによって同じドメインのバイアスが異なる効果を及ぼす可能性が高いため、実質的なバイアス mate-

✉: 小島原典子 〒162-8666 東京都新宿区河田町 8-1 東京女子医科大学医学部衛生学公衆衛生学（二）講座

E-mail: kojimahara.noriko@twmu.ac.jp

doi: 10.24488/jtwmu.88.Extra1_E38

Copyright © 2018 Society of Tokyo Women's Medical University

rial bias を正確に知ることは困難である。リスクとしてバイアスを評価せざるを得ないため、バイアスリスク risk of bias と呼び、確率論的にバイアスを評価することが行われる。

バイアスのモデル化については、Turner RM ら⁶⁾の提唱する方法が知られており、インフルエンザワクチンの効果に関するメタアナリシスに用いた報告がある⁷⁾。このバイアスのモデルでは、バイアスの効果の大きさをリスク比などの効果指標の値、その不確実性を分布と分散で表し、実際に研究で得られた効果指標の値を調整した上で、メタアナリシスを行うことができる。さらに、非直接性は外的バイアスとして評価し、同じ方法で、効果指標の値の調整に用いる。したがって、Turner RM らが内的バイアスおよび外的バイアスと呼ぶバイアスリスクと非直接性を、不確実性を伴った、効果指標の値として導出し、これらの影響を受けた効果指標の値を調整した統合値が得られる。一方で、Turner RM らの提唱する方法は、バイアスのドメインごとに評価した上で、全体としての効果指標の値への影響を調整する精密な方法であり、著者がすでに報告⁸⁾したようにExcelなどを用いて分類・整理しながらまとめることができるが、煩雑なため、広く普及するには至っていない⁹⁾。本論文では、Turner RM らのバイアスのモデルに基づいて、一つの研究の非直接性を含めたバイアス全体の効果をリスク比として表し、その不確実性は標準偏差あるいは95%信頼区間で表し、バイアスの効果で調整したメタアナリシスを行うことで、バイアスを可視化することを試みた。

方 法

1. バイアスのモデル

研究*i*において、*J*個のバイアスが認められた場合のモデルでは、以下の式で表すことができる。実際のRプログラムと解説はAppendixを参照してほしい。

J biases, $j = 1, \dots, J$

δ_{ij}^b : bias j of study i

$\delta_{ij}^b \sim N(\mu_{ij}^b, \sigma_{ij}^{b2})$

$\delta_i^b \sim N(\mu_i^b, \sigma_i^{b2})$: $\mu_i^b = \sum_j \mu_{ij}^b$ $\sigma_i^{b2} = \sum_j \sigma_{ij}^{b2}$

$\delta_i^b \sim N(\sum_j \mu_{ij}^b, \sum_j \sigma_{ij}^{b2})$

$y_i \sim N(\theta_i + \mu_i^b, s_i^2 + \sigma_i^{b2})$

$y_i - \mu_i^b \sim N(\theta_i, s_i^2 / \alpha_i)$: $\alpha_i = s_i^2 / (s_i^2 + \sigma_i^{b2})$

なお、これらの式はすべて対数変換した上で成立する。Nは正規分布を表し、カッコ内は平均値と分

散を示す。 δ_i^b は研究*i*における*J*個のバイアスのバイアス効果の値の合計値を平均値とし、分散の合計値を分散とする正規分布に従うことになる。 y_i は研究*i*で得られたリスク比の値の自然対数であり、 μ_i^b はその研究のすべてのバイアスの効果のリスク比の自然対数の総計である。バイアスで調整した介入の効果は最後の式で表される。すなわち、各バイアスのリスク比の自然対数を加算した総計 μ_i^b を効果指標の値の自然対数から減算し、分散はバイアスの分散を加算した総計 σ_i^{b2} を効果指標の値の分散に加算することで、バイアスで調整されたそれぞれの値が得られる。

以上述べたバイアスはリスク比の乗算、自然対数に変換した後には加算される効果なので、Turner RM らは加算バイアスと呼んでいる。それに対して、対数スケール上でリスク比の値を乗算するとその効果の結果が表されるようなバイアスを比例バイアスと呼んでいる。今回の検討では、実行の容易さを優先し、さまざまなドメインのバイアス全体としての効果の一つにまとめ、リスク比とその標準偏差（対数スケール上）あるいは95%信頼区間で表し、加算バイアスとしてリスク比の値を調整した。

バイアスの効果の大きさは、“もし介入には効果がない場合、サンプリングエラーはないとして、このバイアスだけでどれくらいの見かけ上の効果が表れるか？”に答えるつもりで、それをリスク比で表した値を設定する。非直接性の評価の際には、非直接性が効果推定値に及ぼす効果をバイアスとしてとらえ、“もし介入には効果がない場合、バイアスがない理想の研究であっても、この外的バイアスだけでどれくらいの見かけ上の効果が表れるか？”に答えるつもりで同じように設定する。

Turner RM ら¹⁰⁾はバイアス判定の目安としてバイアスなし (none) : リスク比 1.0, バイアス低 (low) : リスク比 0.9~1.0, バイアス中 (medium) : リスク比 0.7~0.9, バイアス高 (high) : リスク比 < 0.7 という基準を提示している。リスクが上昇する方向のバイアスの場合はこれらの数値の逆数となる。バイアスの効果の不確実性は、これらリスク比に対する標準偏差あるいは信頼区間で表される。メタアナリシスはRパッケージのmetaforとForest plot作成のためのパッケージであるforestplotを用いた。PC環境としてはWindows10またはMac OSXいずれでも動作可能であることを確認済みである。

Table 1 Data prepared on Excel for analysis

author	nc	rc	nt	rt	label	bias	biasusd
Salminen P 2015	273	272	257	186	Author Year	1	1.1
Vons C 2011	119	104	120	72	Surgery	1	1.1
Hansson J 2009	167	142	202	83	Antibiotics	0.65	1
Styrud J 2006	124	124	128	97	Healing	0.8	0.95
Eriksson S 1995	20	20	20	13	Risk Ratio RR	0.8	0.95

Author: Research ID; nc: total number of patients in the control group; rc: number of risk events in the control group, nt: total number of patients in the treatment group; rt: number of risk (event) in the treatment group label: Research ID to be displayed in a forest plot; bias: the degree of indirectness and risk that is expressed as a risk ratio; biasusd: the value of +1 standard deviation, the value of risk ratio before logarithmic transformation.

The value obtained by subtracting the logarithm of the bias value from that of the biasusd value corresponds to one standard deviation of the normal distribution on a logarithmic scale. The value of -1 standard deviation is the exponential of the value obtained by subtracting that value from the logarithm of the bias value, which is calculated from the values of bias and biasusd. Furthermore, the probability of falling within the range of ± 1 standard deviation is twice that of falling within any other range. On the other hand, when using the upper limit value of the 95% confidence interval, the label used is biasuci. The pre-logarithmic-transformation value of the risk ratio is used in this case as well. Subtracting the logarithm of biasuci from the logarithm of bias and dividing the resulting value by 1.96 yields a value equivalent to one standard deviation of the normal distribution on a logarithmic scale. The difference between the value of bias and the values of both biasusd and biasuci could be interpreted as an index of the uncertainty of the bias effect.

2. 解析用データの準備

メタアナリシスのためのサンプルデータを Table 1 に示す。成人の急性虫垂炎で手術を対照とし、抗菌薬投与による保存的治療を介入とした5件のRCTのデータである。アウトカムは“再発を伴わない1年以内の治癒”である。このデータの例におけるバイアスの評価は、最近の2つの研究 Salminen P¹¹⁾、Vons C¹²⁾はバイアス、非直接性は低、方向性は不明、Hansson J¹³⁾はバイアス、非直接性が高く介入群の治癒率が低い方向へ過大評価されており（抗菌薬の効果としては過小評価）、その程度も大きくまた不確実性も高い、Styrud J¹⁴⁾、Eriksson S¹⁵⁾は治癒率が低い方向へ過大評価されているが、その程度は中程度で、不確実性は低いと判定した場合である（Table 2）。

結果および考察

Fig. 1 に示すように、ランダム効果モデルによる通常のメタアナリシスの統合値はリスク比 0.662 で、95%信頼区間上限値 0.778 となった。研究間の異質性の統計学的指標のひとつである I^2 値は 85.7% と異質性が高いことを示唆している。Hansson J は症例減少バイアスが大きい研究であり、真の値からかなりのずれが推定される。ランダム効果モデルを

用いているため、サンプルサイズの小さな研究の Weight は低くなるが、統合値からのずれが大きい研究も Weight が低くなる。したがって、Styrud J と比べると、Hansson J はサンプルサイズが大きいにも関わらず、Weight は小さくなっており、ランダム効果モデルの特性が表れている。

バイアス調整後には、過大評価されており、不確実性も大きいと判断した研究 Hansson J は効果指標の値が右側へずれ、1により近づいた値になっているとともに、信頼区間の幅が拡大して、不確実性が高いという評価が反映されている（Fig. 2）。そのため、Weight も 19.6 から 2.4% に低下している。興味深いことに、バイアス調整後の統合値は 0.751 と、より介入の効果が高い方に移動しており、上限値も 0.859 と、より大きな値になっているとともに、95%信頼区間の幅は狭くなっている。Hansson J の値が上記のごとく調整された結果および各研究の分散が大きくなった効果以上に各研究の効果指標の値が調整された結果研究間のばらつきが小さくなったためと考えられる。

バイアス調整前の Forest plot に調整後の各研究の効果指標、統合値とそれらの信頼区間を図として

Table 2 Results on evaluation of indirectness and risk of bias among five randomized controlled trials (low (0); unclear/moderate (-1); high (-2))

個別研究	バイアスリスク*										非直接性*				リスク人数 (アウトカム率)				効果指標 (種別)	効果指標 (値)	信頼区間
	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	不完全アウトカムデータ	選択的アウトカム報告	早期試験の中止	その他	まとめ	対照群/分子	対照群/分子	介入群/分子	介入群/分子	介入群/分子	介入群/分子	介入群/分子	介入群/分子				
研究コード	研究デザイン	ランダム化	シリアルメソッド	コンソール	盲検化	盲検化	盲検化ITT	不完全アウトカムデータ	選択的アウトカム報告	早期試験の中止	その他	まとめ	対照群/分子	対照群/分子	介入群/分子	介入群/分子	介入群/分子	効果指標 (種別)	効果指標 (値)	信頼区間	
Salminen P 2015	RCT	-1	-1	-1	-1	-1	0	0	0	0	0	-1	0	-1	0	0	0	RR	0.726	0.67 ~ 0.78	
Vons C 2011	RCT	-1	-1	-1	-1	-2	0	0	0	0	-1	-1	0	-1	0	0	0	RR	0.69	0.58 ~ 0.81	
Hansson J 2009	RCT	-2	-1	-1	-1	-2	0	0	0	0	0	-2	-2	-1	0	0	0	RR	0.48	0.40 ~ 0.58	
Styrud J 2006	RCT	-1	0	-1	-1	0	0	0	0	0	-1	-2	-1	0	0	0	0	RR	0.76	0.69 ~ 0.84	
Eriksson S 1995	RCT	-1	-1	-1	-1	0	0	0	0	0	-1	-2	-1	0	0	0	0	RR	0.66	0.48 ~ 0.91	
計													703	662	727	451					

重ねて表示したのが Fig. 3 である。なお、個別研究の効果指標の値と 95 % 信頼区間の上限値・下限値の値は研究間のバラツキの指標である τ^2 が加味されていない通常のメタアナリシスの場合 (Fig. 1) と同じ結果を示す。また、Table 1 に示すバイアスの効果を Forest plot の形式で表したのが Fig. 4、これらのバイアスの効果で Fig. 1 に示す効果指標を調整したものが Fig. 2 といえる。

バイアスリスクおよび非直接性の効果の大きさも方向性も判定せず、効果推定値の不確実性を大きくする効果だけを判定し、それをメタアナリシスの結果に反映させた場合にどうなるか見てみよう。Fig. 5 にバイアスの効果を可視化して Forest plot として表示し、Fig. 6 にバイアス調整前後のメタアナリシスの結果を示す。Turner RM らのモデルでは、リスク比の分散にバイアスの分散が加算されるため通常バイアス調整後のメタアナリシスでは統合値の信頼区間が拡大するケースが多いはずであるが、この例では逆に狭くなった。Hansson J の重みが 2.8 % と無視しうるほどに減少したため、研究間のばらつきが小さくなったためと考えられる。一方、感度分析としてバイアスリスクも非直接性も大きい Hansson J を除いて、その他の条件は Fig. 6 と同じにしてバイアス調整メタアナリシスを行うと、Fig. 7 に示すように統合値の信頼区間は拡大した。

RCT におけるバイアスリスクの評価法については、コクランレビューで用いられている方法がよく知られている¹⁶⁾¹⁷⁾。Cochrane risk of bias tool と呼ばれているが、現在 Version 2.0 の草案が公開されている¹⁸⁾。この新バージョンでは、例えば、今まで Other biases に分類されていた Baseline imbalances を Bias arising from randomization process すなわち選択バイアスに含め、“Was the allocation sequence random?”、“Was the allocation sequence concealed?”、“Were there baseline imbalances that suggest a problem with randomization?” の 3 つのシグナリングクエスチョンに対して、Yes, Probably yes, Probably no, No, No information で答えることで、“High”, “Some concerns”, “Low” の 3 段階で評価するようになっている。

Cochrane Handbook for Systematic Reviews of Interventions も現在の Version 5.1 から 5.2 へ改訂作業中で一部は作業が終了しているようであるが、大きな変更は行われないようである。現在のハンドブックの 8 章 8.5.3 The judgement の項には以下の

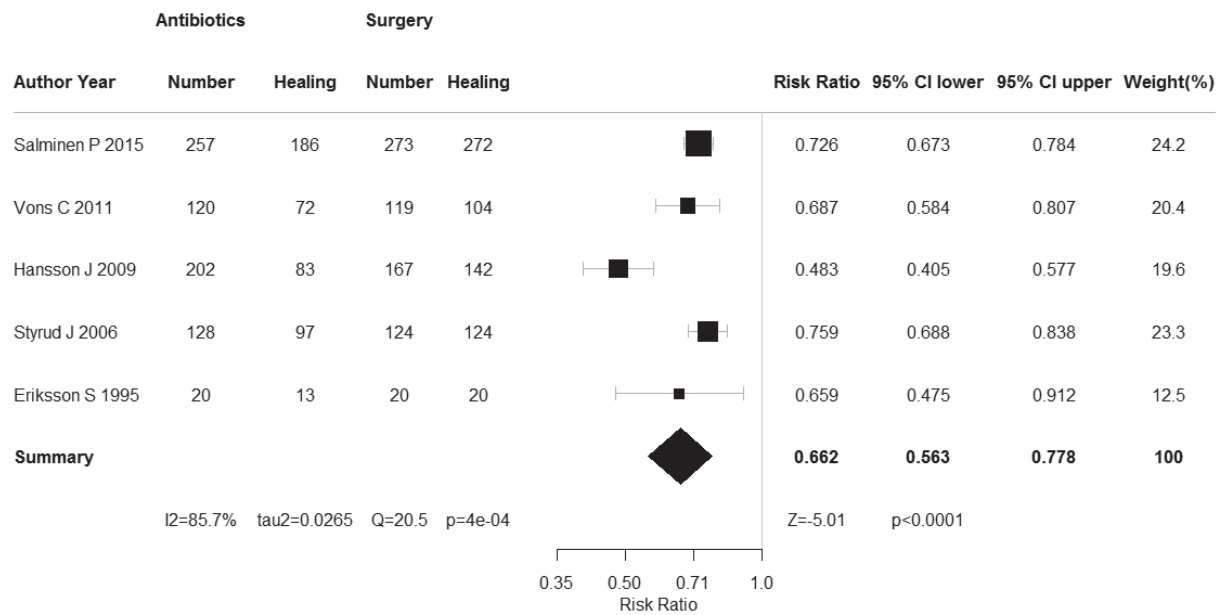


Fig. 1 Forest plots of meta-analysis based on the random-effects model

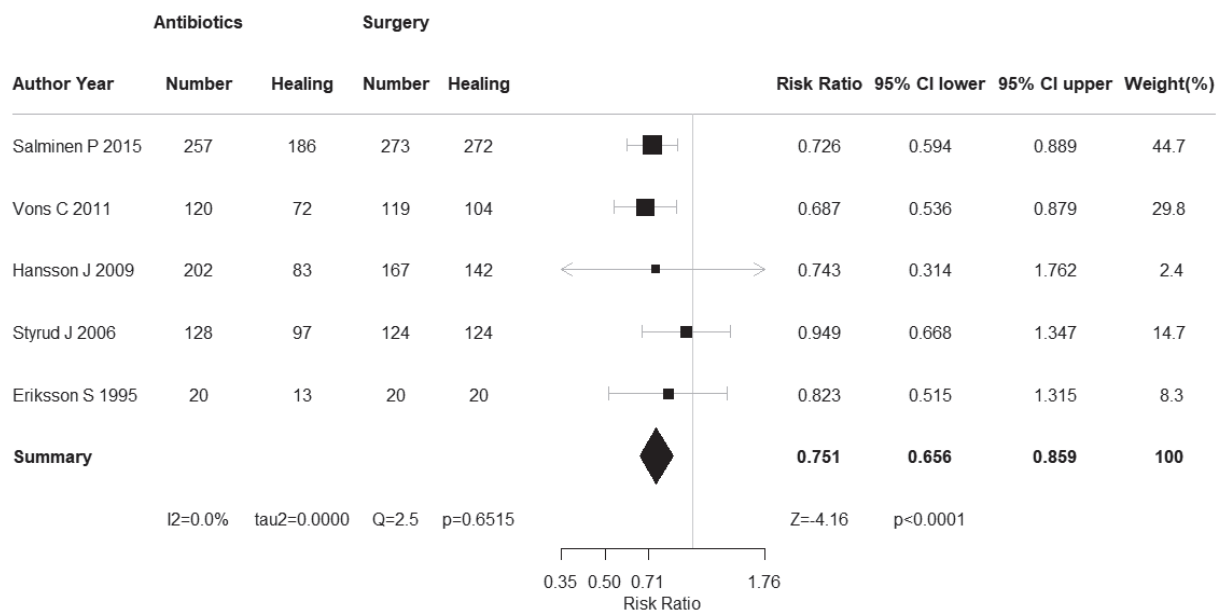


Fig. 2 Forest plots of the meta-analysis after bias adjustment

ように記述されている。“material bias”（リスクとして評価される実質的バイアスという意味に解釈される）を“研究の結果あるいは結論に顕著な影響を持つほど十分な大きさのバイアス”と定義した上で、何らかのバイアスのリスクではなく“material bias”のリスクを考慮すべきであり、そのような判断にはいずれの場合も主観性が含まれる。

バイアスリスク評価の際の評価者間の一致率についても報告があり¹⁹⁾²⁰⁾、その判定には評価者の専門

性、経験などが影響し、バラつきが生じることが示されている。上記 Cochrane handbook でも臨床疫学、医学統計学、システマティックレビューのエキスパートとコンテンツのエキスパートすなわち疾患の専門家がペアで研究の評価を行うことが推奨されている。たとえば、Baseline imbalances の一つとして、比較される群間で年齢分布が異なっていたり、重症度が異なっていたりした場合に、それがアウトカムにどのように影響するのかを推定するのは疾患

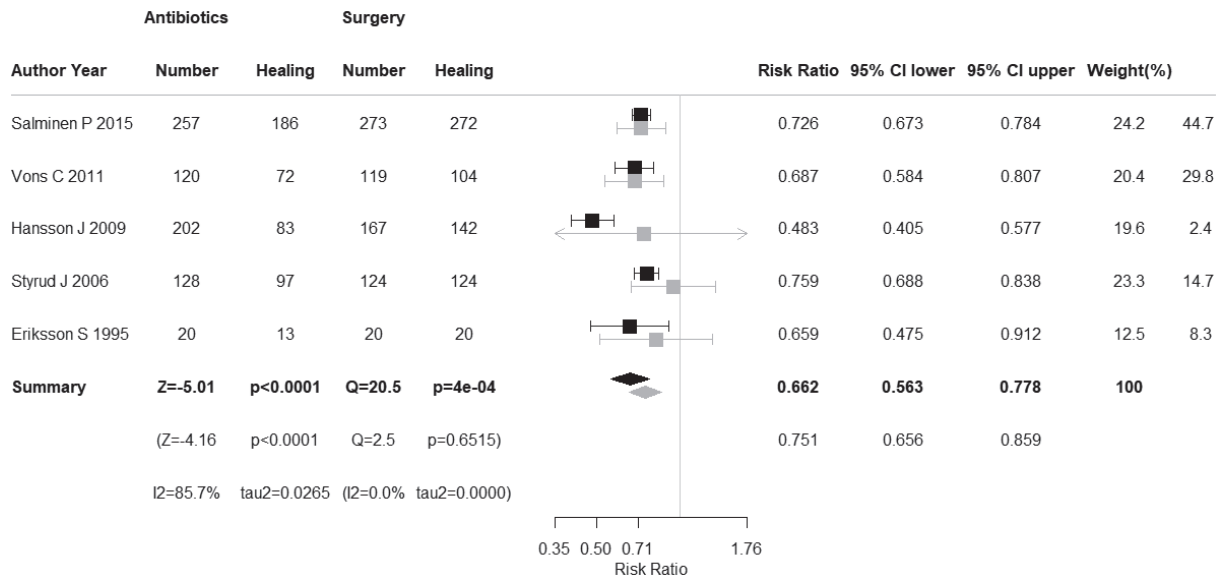


Fig. 3 Forest plots of the meta-analysis before and after bias adjustment

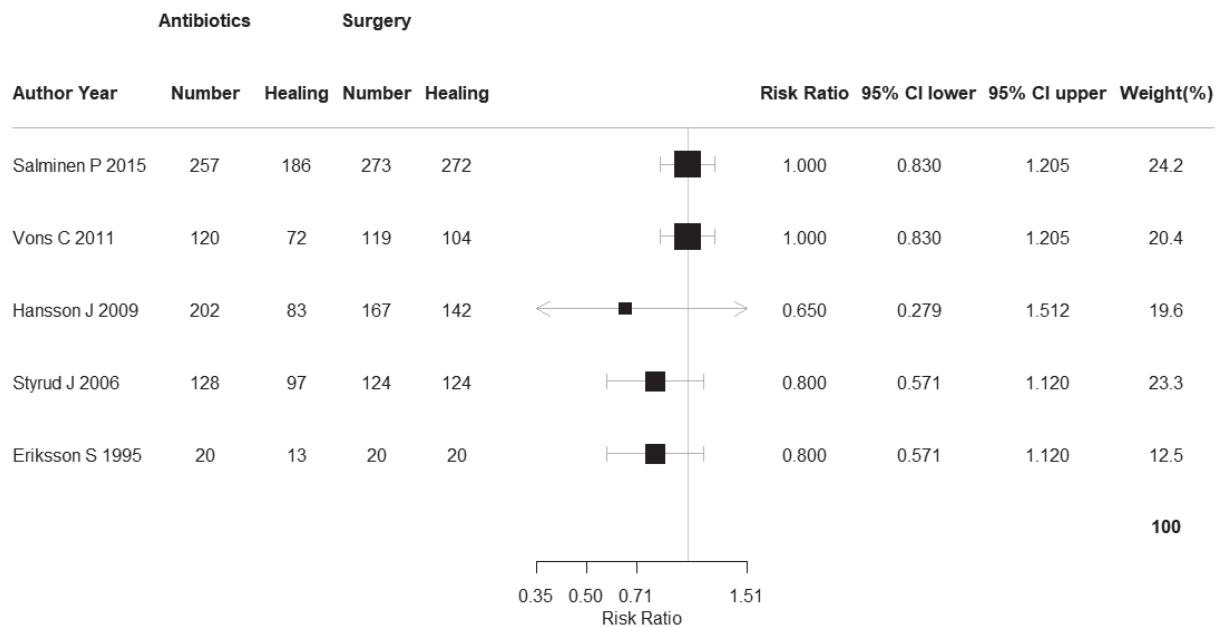


Fig. 4 The bias effect expressed as risk ratio and 95 % confidence interval limits

専門家には容易であっても、システマティックレビューの専門家には困難である。

また、非直接性の評価は特に疾患専門家以外には困難であり、日本人の患者に適用した場合の効果を推定する場合には、疾患専門家でなければ困難と考えられる。バイアスリスクの評価の際にバイアスの大きさ magnitude と方向 directions, すなわち過大評価か過小評価か、について評価することは Cochrane handbook には推奨されるレベルであり必須ではないと述べられているため、バイアスリス

クの各ドメインを High, Unclear, Low の3分類で評価するにとどまり、例えば High と判定された場合に、研究結果に顕著な影響を及ぼしているバイアスの可能性が高いという以上の情報は提供されない。もし2つの異なるドメインのバイアスの方向性が異なる場合には、その研究結果の効果推定値は真の値からずれていない可能性が高くなる。しかし、High と判定されたドメインが2つあれば、全体として、その研究のバイアスリスクは High と判定されてしまうことになる。その場合の解釈として効果推

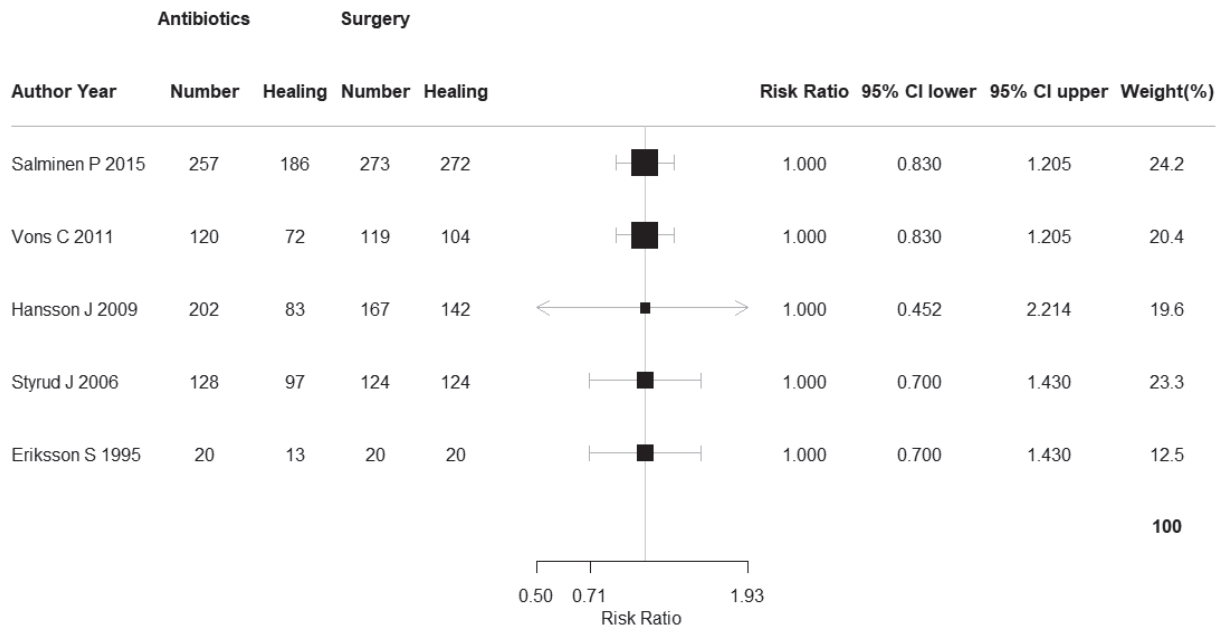


Fig. 5 Visualized bias effect

When only the effect on the uncertainty to the estimated value was assessed, without evaluating the direction and magnitude of the bias, we visualized the bias effect expressed by the risk ratio and the value of +1 SD by using a forest plot.

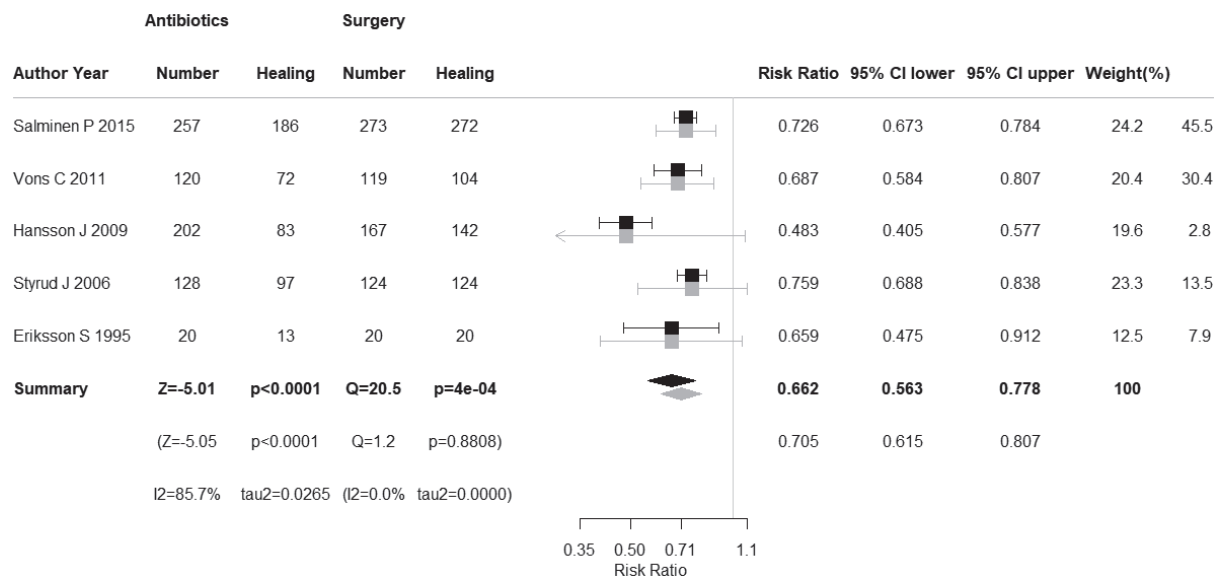


Fig. 6 Forest plots of meta-analysis before and after bias adjustment

If the bias effect is perceived only as an effect on uncertainty, the value of the effectiveness indicators of each study will not change, and the confidence interval will expand. The grey colour indicates having been bias-adjusted, and in this example, based on the use of the random effect model, the integration value approaches 1.0 and the confidence interval narrows.

定値はバイアスにより真の値から大きくずれが生じていると考える者が多いであろう。しかし、効果推定値は変動せず、不確実性のみが大きくなるはずである。

本稿で示した方法を用いることによって、①バイアスの効果の大きさと方向性が推定できる場合（リスク比の値が設定でき分布の大部分が>1.0または<1.0にある場合）も、②大きさは推定できるが方向

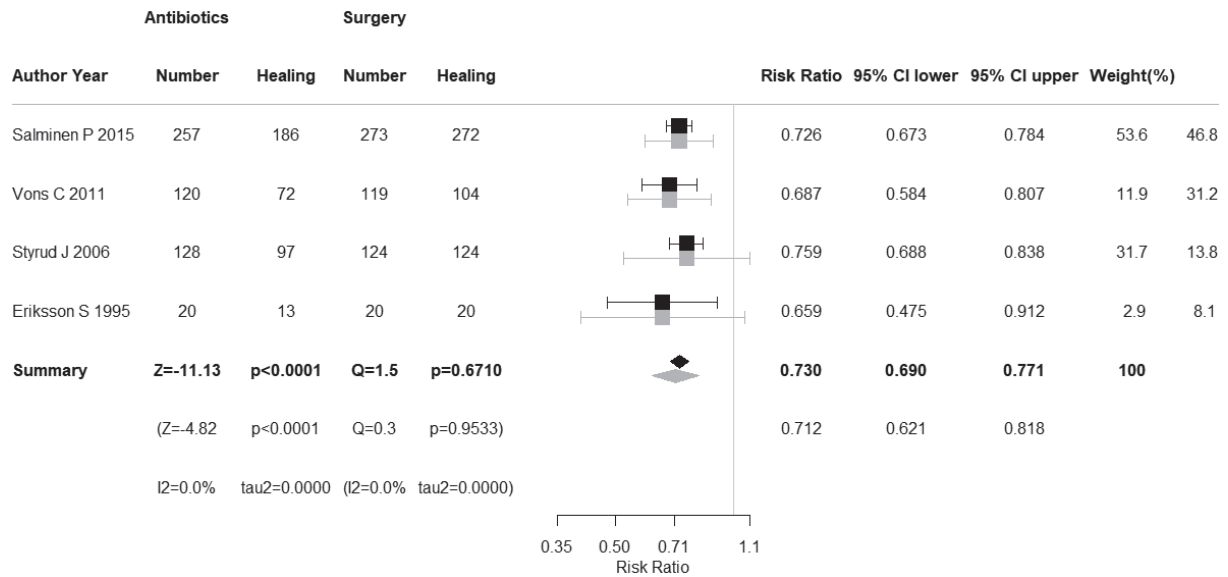


Fig. 7 Results of the sensitivity analysis of the meta-analysis

Excluding Hansson J 2009, when only the effect on the uncertainty of each study was evaluated, the integrated value decreased from 0.730 (95 %CI 0.690-0.771) to 0.712 (0.621-0.818) following bias adjustment. Thus, the confidence interval expanded and uncertainty increased.

性が推定できない場合（リスク比の値が設定できるが分布が1.0をまたぐ場合）も、③大きさも方向性も推定はできないがバイアスのリスクが高いと判定できる場合（リスク比の値を1.0に設定しバイアスのリスクの程度に応じて分布の広がりを設定する場合）も、いずれの場合も調整した結果に反映させることができ、メタアナリシスで統合した場合の影響も容易に知ることができる。さらに、感度分析としてさまざまなバイアスの大きさ、方向性、不確実性を設定して、統合値がどのように変動するかを見ることによって、エビデンス総体の確実性を検討することが可能となる。

結 論

Cochraneの方法のように、方向性、大きさについて明確な判断をしない場合は、High risk of biasの判定に対してバイアスの効果の標準偏差を大きく、Low risk of biasの判定に対しては小さく設定してバイアス調整メタアナリシスを実行することにより、バイアスが可視化され、エビデンスの確実性の評価が容易になる可能性がある。バイアスの効果の標準偏差をどの程度に設定するかは評価者の判断により、バイアスの効果の不確実性に応じて設定する。また、バイアスの評価を複数の評価者が行うことも可能であり、その場合、評価ドメイン、項目、バイアスの効果の判定方法について、あらかじめ議論し、

理解を共有した上で、独立して評価を行う。結果について、議論した上で思い違いや判定エラーがないか確認し、大きさと標準偏差について中央値を用いて、バイアス調整メタアナリシスを実行することが可能である。評価者が多数の場合、大きさと標準偏差の分布を示したり、中央値と最小値、最大値を示したりすることで、バイアスリスクの評価に透明性を確保することが可能になる。また、従来行われている方法の様に、2名の評価者で意見を調整して一つの評価にまとめる必要がなくなる。

今後、現在一般的に行われている定性的なバイアスリスクの評価法と、本稿で示したバイアスの大きさ、方向、不確実性の3つの要素を定量的に評価する方法とを比較する研究が必要であろう。その際に、本稿で述べた方法は有力なツールになると考える。

利益相反は特にない。

文 献

- 1) IOM (Institute of Medicine). Clinical Practice Guidelines We Can Trust. The National Academies Press, Washington, DC (2011) <http://www.nationalacademies.org/hmd/Reports/2011/Clinical-Practice-Guidelines-We-Can-Trust.aspx> (accessed on Nov 23, 2017)
- 2) IOM (Institute of Medicine). Standards for Systematic Reviews. The National Academies Press, Washington, DC (2011) <http://www.nationalacademies.org/hmd/Reports/2011/Standards-for-Systematic-Reviews.aspx>

- nationalacademies.org/hmd/Reports/2011/Finding-What-Works-in-Health-Care-Standards-for-Systematic-Reviews/Standards.aspx (accessed on Nov 23, 2017)
- 3) **Atkins D, Best D, Briss PA et al: Grade Working Group:** Grading quality of evidence and strength of recommendations. *BMJ* **328**: 1490, 2004
 - 4) **GRADE handbook** 2013. <http://gdt.guidelinedevelopment.org/app/handbook/handbook.html#h.svwngs6pm0f2> (accessed on Nov 23, 2017)
 - 5) **Berkman ND, Santaguida PL, Viswanathan M et al:** The Empirical Evidence of Bias in Trials Measuring Treatment Differences. *Methods Research Report*. Agency for Healthcare Research and Quality, Rockville (2014)
 - 6) **Turner RM, Spiegelhalter DJ, Smith GC et al:** Bias modelling in evidence synthesis. *J R Stat Soc Ser A Stat Soc* **172**: 21–47, 2009
 - 7) **Darvishian M, Gefenaite G, Turner RM et al:** After adjusting for bias in meta-analysis seasonal influenza vaccine remains effective in community-dwelling elderly. *J Clin Epidemiol* **67**: 734–744, 2014
 - 8) **森實敏夫**：医学統計学シリーズ第35回 バイアスによる効果指標の調整. *あいみっく* **37**(1)：8–18, 2016
 - 9) **Dias S, Sutton AJ, Welton NJ et al:** Evidence synthesis for decision making 3: heterogeneity-subgroups, meta-regression, bias, and bias-adjustment. *Med Decis Making* **33**: 618–640, 2013
 - 10) **森實敏夫**：医学統計学シリーズ第40回 バイアスで調整したメタアナリシス. *あいみっく* **38**(1)：10–19, 2017
 - 11) **Salminen P, Paajanen H, Rautio T et al:** Antibiotic therapy vs appendectomy for treatment of uncomplicated acute appendicitis: The APPAC randomized clinical trial. *JAMA* **313**: 2340–2348, 2015
 - 12) **Vons C, Barry C, Maitre S et al:** Amoxicillin plus clavulanic acid versus appendectomy for treatment of acute uncomplicated appendicitis: an open-label, non-inferiority, randomised controlled trial. *Lancet* **377**: 1573–1579, 2011
 - 13) **Hansson J, Körner U, Khorram-Manesh A et al:** Randomized clinical trial of antibiotic therapy versus appendectomy as primary treatment of acute appendicitis in unselected patients. *Br J Surg* **96**: 473–481, 2009
 - 14) **Styrud J, Eriksson S, Nilsson I et al:** Appendectomy versus antibiotic treatment in acute appendicitis. a prospective multicenter randomized controlled trial. *World J Surg* **30**: 1033–1037, 2006
 - 15) **Eriksson S, Granström L:** Randomized controlled trial of appendectomy versus antibiotic therapy for acute appendicitis. *Br J Surg* **82**: 166–169, 1995
 - 16) **Higgins JP, Altman DG, Gøtzsche PC et al: Cochrane Bias Methods Group: Cochrane Statistical Methods Group:** The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* **343**: d5928, 2011
 - 17) **Cochrane handbook for systematic reviews of interventions (version 5.1.0).** <http://training.cochrane.org/handbook> (accessed on Nov 23, 2017)
 - 18) **Higgins JPT, Savović J, Page MJ, Sterne JAC, ed:** Revised Cochrane risk of bias tool for randomized trials (RoB 2.0), 2016. (accessed on Nov 23, 2017) https://www.bristol.ac.uk/media-library/sites/social-community-medicine/images/centres/cresyda/RoB2-0_indiv_main_guidance.pdf
 - 19) **Armijo-Olivo S, Ospina M, da Costa BR et al:** Poor reliability between Cochrane reviewers and blinded external reviewers when applying the Cochrane risk of bias tool in physical therapy trials. *PLoS One* **9**: e96920, 2014
 - 20) **Hartling L, Hamm M, Milne A et al:** *In* Validity and Inter-Rater Reliability Testing of Quality Assessment Instruments. Agency for Healthcare Research and Quality (US), Rockville (2013)