

第28回日本疫学会学術総会

JE編集委員会企画

疫学研究に求められる観察研究の報告ガイドライン

観察研究のメタアナリシス ～報告ガイドライン～

2018年2月1日

国立がん研究センター

社会と健康研究センター疫学研究部

後藤温

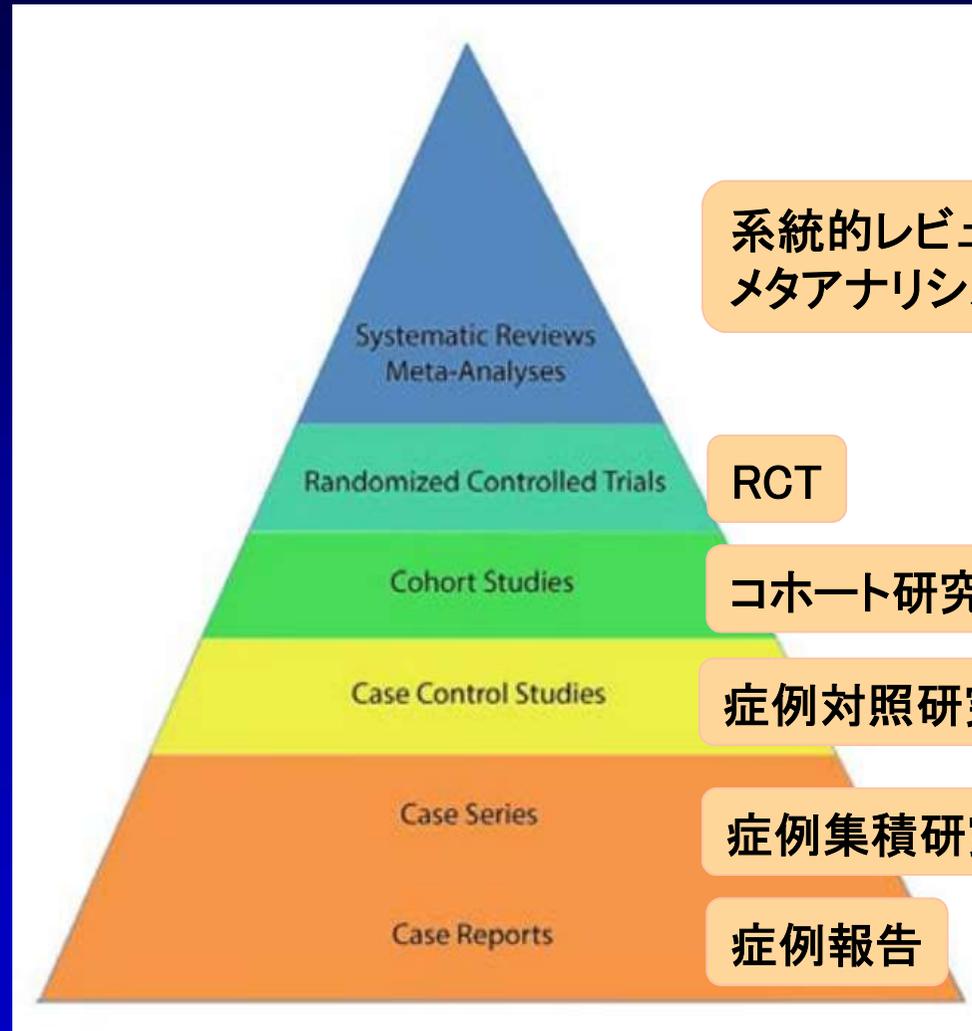


国立研究開発法人

国立がん研究センター

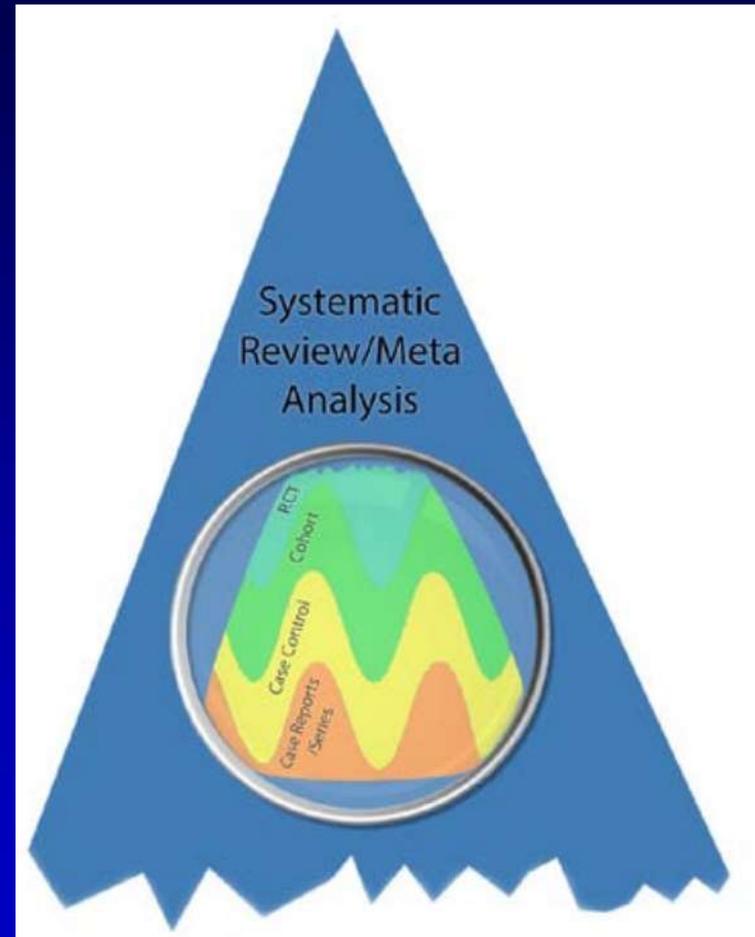
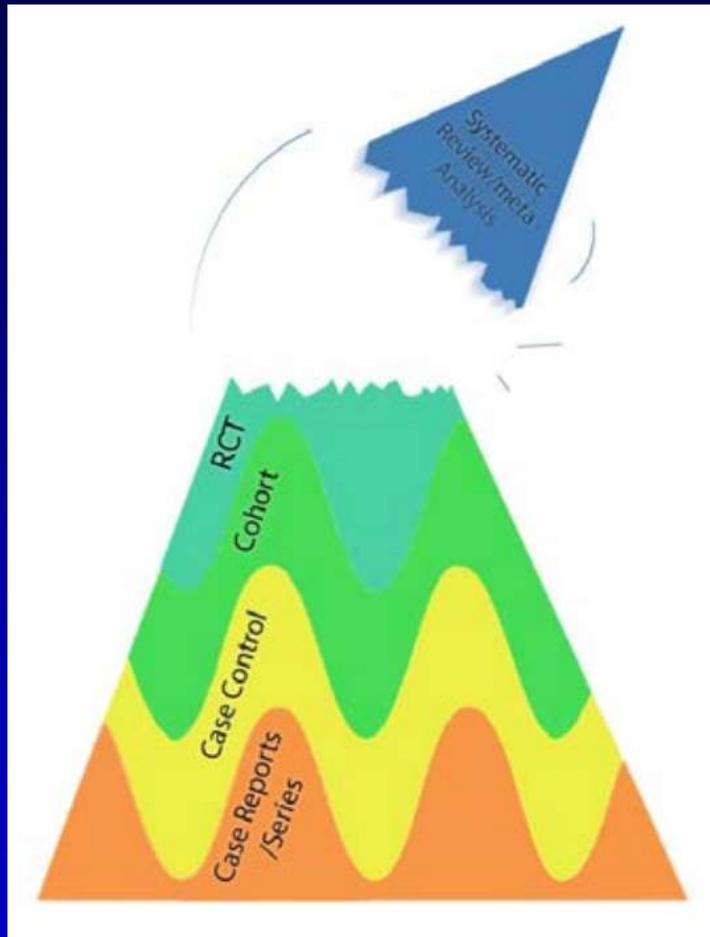
National Cancer Center Japan

(従来の)エビデンスのピラミッド



実際は、エビデンスのピラミッド(ヒエラルキー)は絶対的なものではない

エビデンスの(新)ピラミッド



喫煙と肺がんのように、観察研究がRCTよりも強力なエビデンスを提供することがある
「最善」のエビデンスを得るには系統的レビュー・メタアナリシスがほぼ必須

系統的レビュー/メタアナリシス (SR/MA)

系統的レビュー (Systematic Review: SR) とは

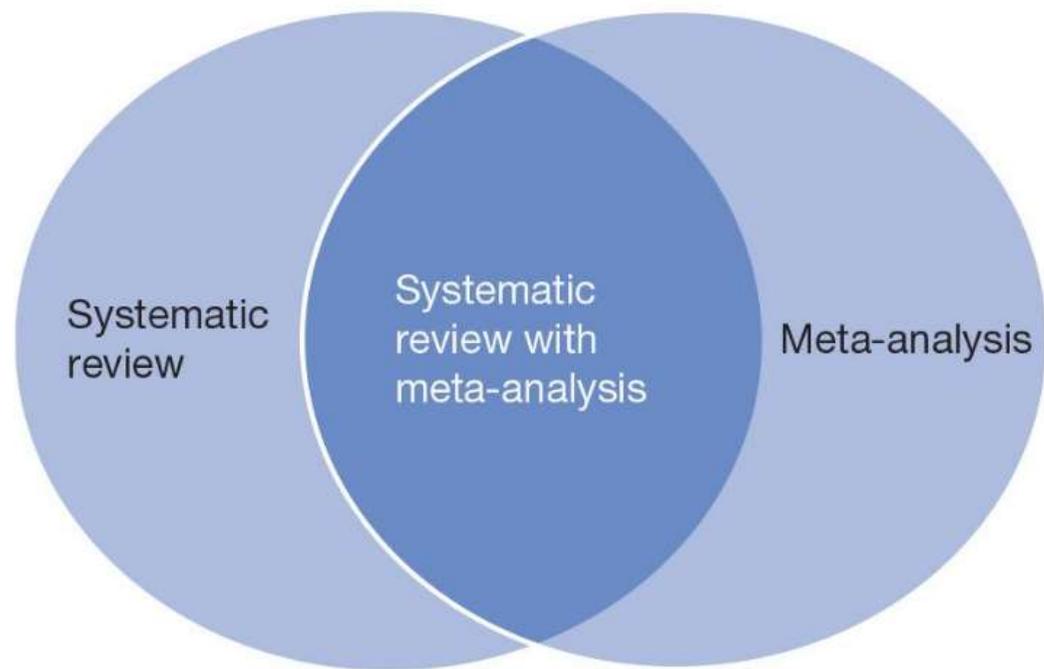
明確な研究目的に対して、再現性のある方法で、
適格基準を満たす全研究把握のために網羅的検索を行い、
研究結果の妥当性を吟味し、
結果や各研究の特徴を系統的に提示し、要約するものである。

メタアナリシス (Meta-analysis: MA) とは

研究結果を統計学的手法で統合し要約するものである。

SR & MA

The Overlap of Study Designs: Systematic Review and Meta-analysis

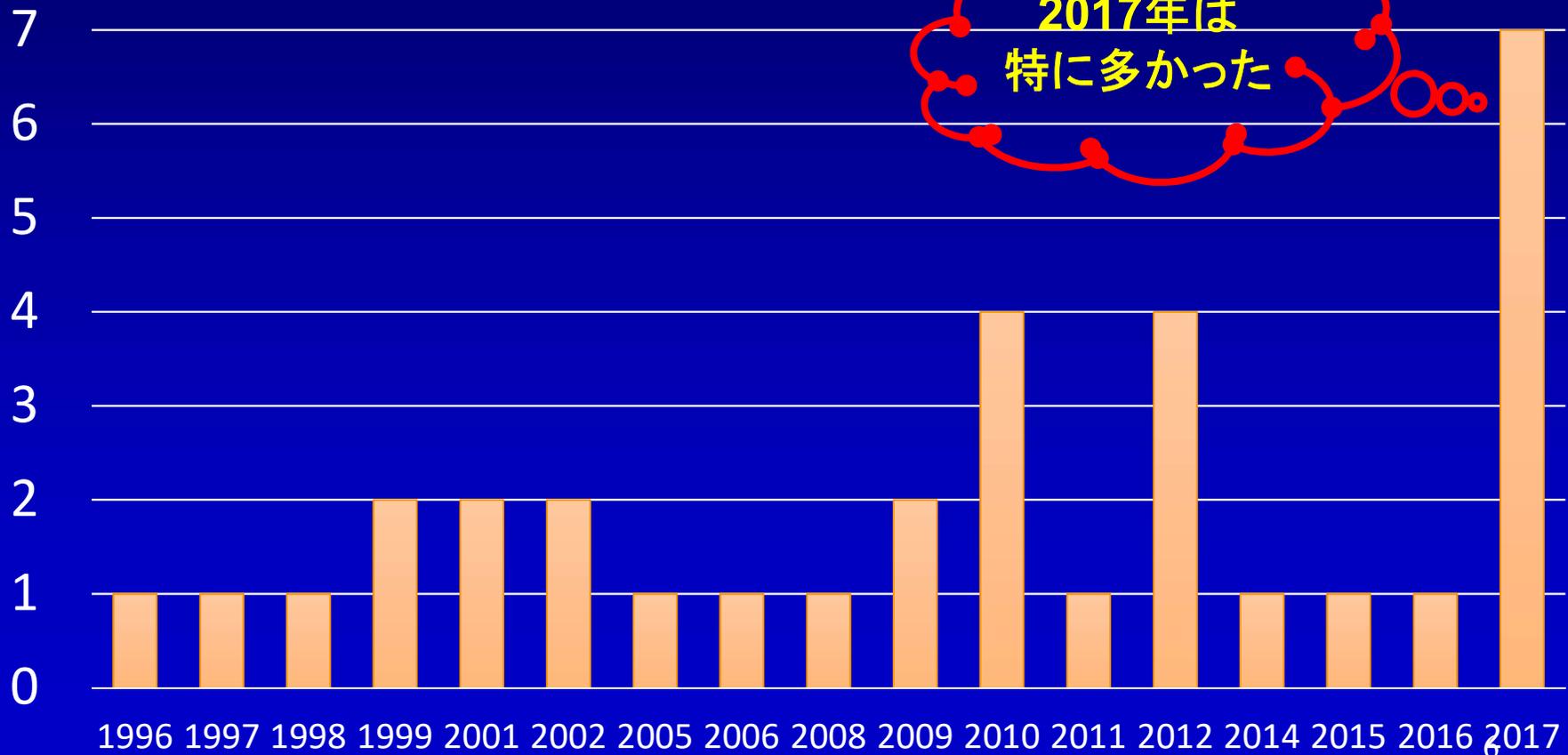


Journal of Epidemiology

SR論文数

PUBMED検索: "J Epidemiol"[jour] AND systematic[sb]

線形トレンドP値=0.13





PRISMA

TRANSPARENT REPORTING OF SYSTEMATIC REVIEWS AND META-ANALYSES

- PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

<http://www.prisma-statement.org>

日本語訳

<http://www.prisma-statement.org/Translations/Translations.aspx>

Meta-analysis of Observational Studies in Epidemiology

A Proposal for Reporting

JAMA, April 19, 2000—Vol 283, No. 15

- MOOSE : Meta-analysis Of Observational Studies in Epidemiology

<https://jamanetwork.com/journals/jama/fullarticle/192614>

日本語訳

http://cont.o.oo7.jp/MOOSE/MOOSE_chk_Jap.pdf

ガイドラインのスコープ



PRISMA

TRANSPARENT REPORTING OF SYSTEMATIC REVIEWS AND META-ANALYSES

著者が透明性の高い、完全な系統的レビューとメタアナリシスを確実に報告できるようにすること

Meta-analysis of Observational Studies in Epidemiology

A Proposal for Reporting

JAMA, April 19, 2000—Vol 283, No. 15

著者、査読者、編集委員、読者が観察研究の系統的レビューとメタアナリシスを適切に報告・評価・活用できるようにすること

PRISMA・MOOSEチェックリストの構成

	PRISMA 項目数	MOOSE 項目数
タイトル	1	0
抄録	1	0
序論/背景	2	6
方法	12	18
結果	7	4
考察	1	3
結論	1	2
資金	1	1
計	27	34

➤ 重要項目は概ね共通している。

➤ PRISMAはRCTを想定した構成であるが、必須事項が網羅的に列挙されており、観察研究でも有用である。

➤ MOOSEは観察研究に特化しており、文献検索の箇所も詳細である。

観察研究SR/MAの必須事項(序論)

- 既知の知見と照らし合わせて、SR/MAの理論的根拠の説明
- 研究デザイン、研究対象、曝露、アウトカム、研究目的の記載

観察研究SR/MAの必須事項(序論)

■ 既知の知見と照らし合わせて、SR/MAの理論的根拠の説明

- なぜこの研究テーマが重要なのか？
- 既知の知見や解決されていない問題はなにか？
- 本SR/MAを実施することで何が分かるのか？

“Given the high prevalence of smoking, ... tobacco control may have a **substantial importance** in managing diabetes in Japan. However, **there has been no systematic evaluation** of the association between smoking and the risk or burden of diabetes in Japan.”

観察研究SR/MAの必須事項(序論)

- 研究デザイン、研究対象、曝露、アウトカム、研究目的の記載

➤ どんなりサーチ・クエスションか？

“Therefore, the present study was performed to provide 1) a quantitative summary of the association between **smoking status** (current smoking, former smoking, smoking cessation years, and passive smoking) and the **risk of T2DM in Japan.**”

観察研究SR/MAの必須事項(方法)

■ プロトコールとその登録

■ 適格基準

■ 情報源と文献検索方法

■ スクリーニング、データ抽出

■ 個々の研究の質の評価法

■ 要約指標と統計手法

観察研究SR/MAの必須事項(方法)

■プロトコールとその登録

- プロトコールを作成し、目的や方法が事前に設定しているか？

“Our **protocol was registered** and is available on PROSPERO, an international prospective register of systematic reviews (registration number: CRD42015027797).”

Hori M, Katanoda K et al. Jpn J Clin Oncol. 2016 Oct;46(10):942-951



The image shows a registration card from PROSPERO. At the top left is the logo for the University of York Centre for Reviews and Dissemination. At the top right is the logo for the NHS National Institute for Health Research. The center text reads "PROSPERO International prospective register of systematic reviews". Below this, the title of the study is "Passive smoking and the risk of lung cancer in Japan: a systematic review and meta-analysis of epidemiologic studies", followed by the authors' names: Kota Katanoda, Megumi Hori, Hirokazu Tanaka.

観察研究SR/MAの必須事項(方法)

■ 適格基準

➤ 明快到適格基準は書かれているか？

“Inclusion criteria included **prospective study design** (including prospective cohort, nested case-control, and case-cohort studies); **exposure of interest** (dietary n-3 PUFAs or blood n-3 PUFAs concentrations); **endpoint** (incident PCa in males); and **reporting of risk estimate** (relative risk, odd ratio, or hazard ratio) of PCa with corresponding **95% confidence intervals** (CIs) for individual n-3 PUFA exposure.”

PRISMA・MOOSEチェックリストの違い

■ 文献検索方法

PRISMA:

少なくとも1つのデータベースの電子検索式について、使用されたあらゆる“limits”を含め、再現できるくらいに詳細に示すこと。

MOOSE : かなり詳細

検索担当者の適格性(例えば、図書館員および調査員)。

検索ストラテジー—統合および検索用語(キーワード)に含まれる期限も含めて。

すべての入手可能な研究を盛り込む努力—著者との連絡も含む。

調査対象となったすべてのデータベースおよび記録。

調査のために使用されたソフトウェアの名前とバージョン—特別機能(explosion など)も含む。

ハンドサーチの使用について(例えば、入手した論文の参考文献リスト)。

特定された文献のリスト および除外された文献のリスト—除外の妥当性に関する理由も含めて。

英語以外で発表された文献の取り扱い方法。

抄録および未発表の研究の取り扱い方法。

いかなる形式であれ著者と連絡をとった事実があれば、それについての説明。

観察研究SR/MAの必須事項(方法)

■ 情報源と文献検索方法

- どの情報源を用いてどの時期の文献を検索したか？
- 検索方法は再現できるように記載されているか？

“We searched two databases (**PubMed** and **EMBASE**) up to **February 2014**. The search strategy was based on the following title/abstract key words: (‘fat’ OR ‘fatty acid’ OR ‘docosahexaenoic acid’ OR ‘eicosapentaenoic acid’ OR ‘docosapentaenoic acid’ OR ‘alpha-linolenic acid’ OR ‘polyunsaturated fatty acid’ OR ‘omega-3 fatty acid’ OR ‘n-3 fatty acid’) AND (‘prostate cancer’ OR ‘prostate neoplasms’)”

Hu FY et al. J Epidemiol. 2015; 25(4): 261–274.

図書館司書や検索専門家にコンサルトするのが望ましい¹⁷

文献検索法に関する査読コメント

“You claim you have performed the study in accordance with MOOSE. However, **you do not completely fulfill the recommendations according to the search strategy** (please recheck the MOOSE statement), e.g., the qualifications of the searcher should be stated. **Is it the investigators, and if yes, were the search controlled with an experienced searcher.**”

図書館司書や検索専門家にコンサルトするのが望ましい

観察研究SR/MAの必須事項(方法)

■スクリーニング、データ抽出

- どのようにして論文をスクリーニングしたか？
- どのようにしてデータを抽出したか？

“We identified articles eligible for further review by performing an initial screen of identified abstracts or titles. The second screening was based on the full-text review. Two investigators (SA and AG) independently assessed the full text for eligibility; discrepancies were resolved via consensus or determined by a third investigator (TM).”

Akter S et al. J Epidemiol. 2017 Dec; 27(12): 553–561.

“For each included study, two authors (S.K. and H.So.) extracted the following information relevant to study characteristics: country, ..., criteria of T2DM cases and non-cases, and used SNPs. Inconsistencies were resolved via discussion.”

Kodama S et al. J Epidemiol. 2018; 28(1): 3–18.

観察研究SR/MAの必須事項(方法)

■ 個々の研究の質の評価法

- 採用された個々の研究の質（バイアス、交絡の可能性）をどのように評価したか？

“Using the **Newcastle-Ottawa Scale**, we assessed the overall quality of each study by totaling scores of the 9 criteria (0–9 stars): the **representativeness of the exposed cohort**, the **selection of the nonexposed cohort**, **ascertainment of exposure**, and **outcome of interest not present at the start of the study** (maximum of 4 stars); **comparability** of the cohorts on the basis of study design and analysis (maximum of 2 stars); and finally, the **assessment of the outcome** (maximum of 3 stars).”

個々の観察研究を評価するツール

Risk of Bias Assessment tool for Non-randomized Studies (RoBANS)

J Clin Epidemiol. 2013 Apr;66(4):408-14

- コクランRCT用Risk of Bias ツールに似ており、評価しやすい
- 福島県立医科大学 竹林 由武 先生の下記スライドをご参照ください

<https://www.slideshare.net/yoshitaket/takebayashi>

Newcastle-Ottawa Scale (NOS)

- 簡便なスケールで一流誌でも頻用されているが、有用性を疑問視する指摘もあり

http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

The ROBINS-E tool (Risk Of Bias In Non-randomized Studies - of Exposures) New!! (preliminary)

- コクランやGRADE等の国際共同チームが開発しているツール(暫定版のみ公開)
- 約50個の質問から構成される

<https://www.bristol.ac.uk/population-health-sciences/centres/cresyda/barr/riskofbias/robins-e/>

個々の観察研究を評価するツール

Newcastle-Ottawa Scale (NOS)

- Selection: 代表性、曝露評価等 (4 items)
- Comparability: 交絡評価 (2 items)
- Outcome (3 items)
の3項目で構成される

曝露評価の質問

- 3) Ascertainment of exposure
 - a) secure record (eg surgical records) ✱
 - b) structured interview ✱
 - c) written self report
 - d) no description

交絡評価の質問

- 1) Comparability of cohorts on the basis of the design or analysis
 - a) study controls for _____ (select the most important factor) ✱
 - b) study controls for any additional factor ✱ (This criteria could be modified to indicate specific control for a second important factor.)

個々の観察研究を評価するツール

The ROBINS-E tool (Risk Of Bias In Non-randomized Studies - of Exposures) New!! (preliminary)

曝露評価の質問

3.1 Is exposure status well defined?	Y / PY / PN / N / NI
3.2 Did entry into the study begin with start of the exposure?	Y / PY / PN / N / NI
3.3 Was information used to define exposure status recorded prior to outcome assessment?	Y / PY / PN / N / NI
3.4 Could classification of exposure status have been affected by knowledge of the outcome or risk of the outcome?	Y / PY / PN / N / NI
3.5 Were exposure assessment methods robust (including methods used to input data)?	Y / PY / PN / N / NI

交絡評価の質問

1.4. Did the authors use an appropriate analysis method that adjusted for all the critically important confounding areas?	NA / Y / PY / PN / N / NI
1.5. If Y or PY to 1.4: Were confounding areas that were adjusted for measured validly and reliably by the variables available in this study?	NA / Y / PY / PN / N / NI
1.6. Did the authors avoid adjusting for post-exposure variables?	NA / Y / PY / PN / N / NI

NI= No Information

観察研究SR/MAの必須事項(方法)

■ 要約指標と統計手法

- どの指標（リスク比、オッズ比等）で評価したか？
- MAに用いた統計手法は何か？

“**Relative risks** (RRs) were used as the common measure of association across studies.”

“We used **DerSimonian and Laird random-effects models** for calculating the summary estimates.”

Akter S et al. J Epidemiol. 2017 Dec; 27(12): 553–561.

“Study **heterogeneity** was assessed using the **I^2 statistic**. An I^2 statistic <25% indicated low heterogeneity, 25%–75% indicated moderate heterogeneity, and >75% indicated high heterogeneity.”

Hu FY et al. J Epidemiol. 2015; 25(4): 261–274. 24

MA: 重み付け平均をすること

名前	点数	重み
山田	95	10
鈴木	90	5
岸田	90	5
後藤	60	1
佐藤	75	1

算術平均(平均値)

$$\frac{95 + 90 + 90 + 60 + 75}{5} = 82$$

重み付け平均

$$\frac{95 \times 10 + 90 \times 5 + 90 \times 5 + 60 \times 1 + 75 \times 1}{10 + 5 + 5 + 1 + 1} = 94.5$$

MAの手順

- Step 1 個々の研究の推定値(ハザード比,信頼区間)を得る
- Step 2 個々の研究の自然対数(e)を底としてlog変換
- Step 3 個々の研究の標準誤差を計算

標準誤差 = (upper limit – lower limit) / 3.92 ※95%信頼区間の時

- Step 4 重み付け平均を計算する

重み付け平均 = $\frac{\log RR \times \text{重みの和}}{\text{重みの和}}$ (RRに戻すには、 $e^{\text{重み付け平均}}$)

重み = $\frac{1}{\text{標準誤差}^2}$

- Step 5 統合値の標準誤差を計算し、95%信頼区間を得る

統合値の標準誤差(SE) = $\frac{1}{\sqrt{\text{重みの和}}}$

95%信頼区間 = 重みづけ平均 $\pm 1.96 \times SE$ (同様にRRスケールに戻す⁶⁾)

MAをやってみよう

Larynx				
Pelucchi, ³³ 2006	Case-control	527		1.23 (0.80–1.90)
Schouten, ¹⁶ 2009	Case-cohort	36		0.93 (0.54–1.58)

logRR	logLL	logUL	SE	SE ²	重み
0.207	-0.223	0.642	0.221	0.049	20.5
-0.073	-0.616	0.457	0.274	0.075	13.3

重みづけ平均=
標準誤差=
95%信頼区間=

MAをやってみよう

logRR	logLL	logUL	SE	SE ²	重み
0.207	-0.223	0.642	0.221	0.049	20.5
-0.073	-0.616	0.457	0.274	0.075	13.3

重みづけ平均= 0.097

標準誤差= 0.172

95%信頼区間= -0.240~0.434

重みづけ平均と信頼区間をeを底としてべき乗すると

RR=1.102 (95%CI, 0.786, 1.543)

MAをやってみよう

実際はStataなどのパッケージを用いるのが便利

Stataコマンド例

```
findit metan
input str12 study year RR lcl ucl
"Pelucchi" 2006 1.23 0.80 1.90
"Schouten" 2009 0.93 0.54 1.58
end
gen SE=(log(ucl)-log(lcl))/(1.96*2)
gen logrr=log(RR)
metan logrr SE, eform
```

MAをやってみよう

実際はStataなどのパッケージを用いるのが便利
Stataコマンド例

```
. metan logrr SE, eform
```

Study	ES	[95% Conf. Interval]	% Weight
1	1.230	0.798 1.896	60.64
2	0.930	0.544 1.591	39.36
I-V pooled ES	1.102	0.787 1.543	100.00

Heterogeneity chi-squared = 0.63 (d.f. = 1) p = 0.427

I-squared (variation in ES attributable to heterogeneity) = 0.0%

Test of ES=1 : z= 0.56 p = 0.573

観察研究SR/MAの必須事項(結果)

■ 研究選択の流れ(フローチャート)

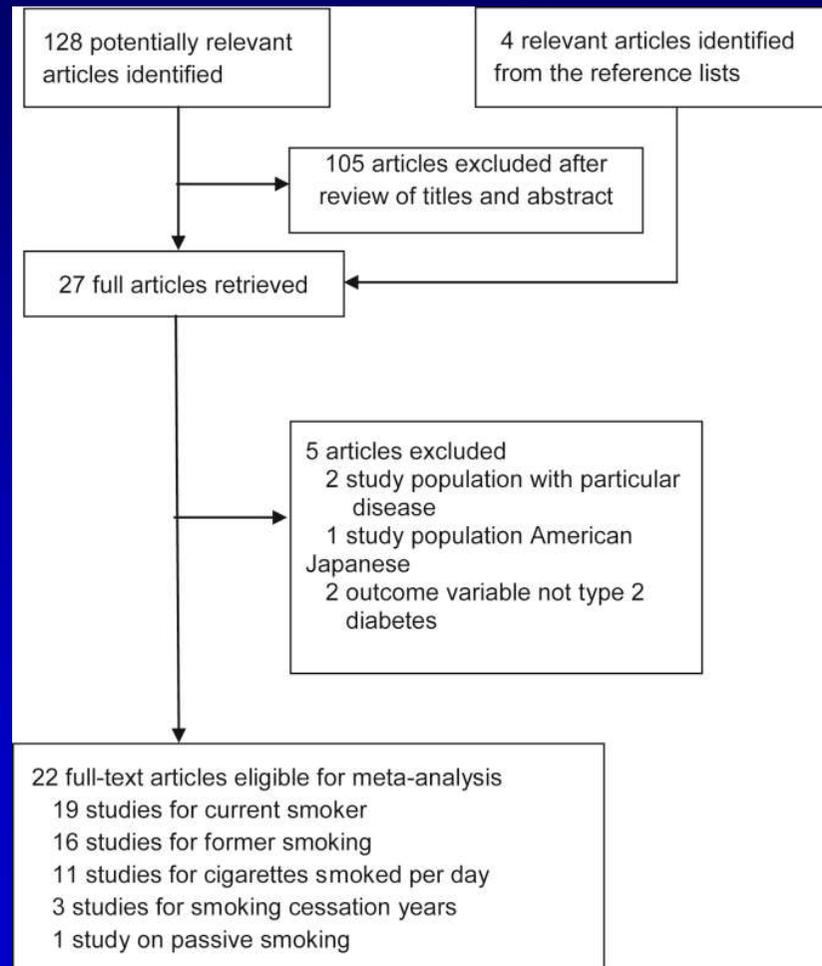
■ 個々の研究の特性や質

■ 個々の研究と研究全体の点推定値と
95%信頼区間等

観察研究SR/MAの必須事項(結果)

■ 研究選択の流れ(フローチャート)

➤ 研究が選択された過程が記載されているか？



観察研究SR/MAの必須事項(結果)

■ 個々の研究の特性や質

- 個々の研究の特性は？
- 個々の研究のバイアスや交絡の可能性は？

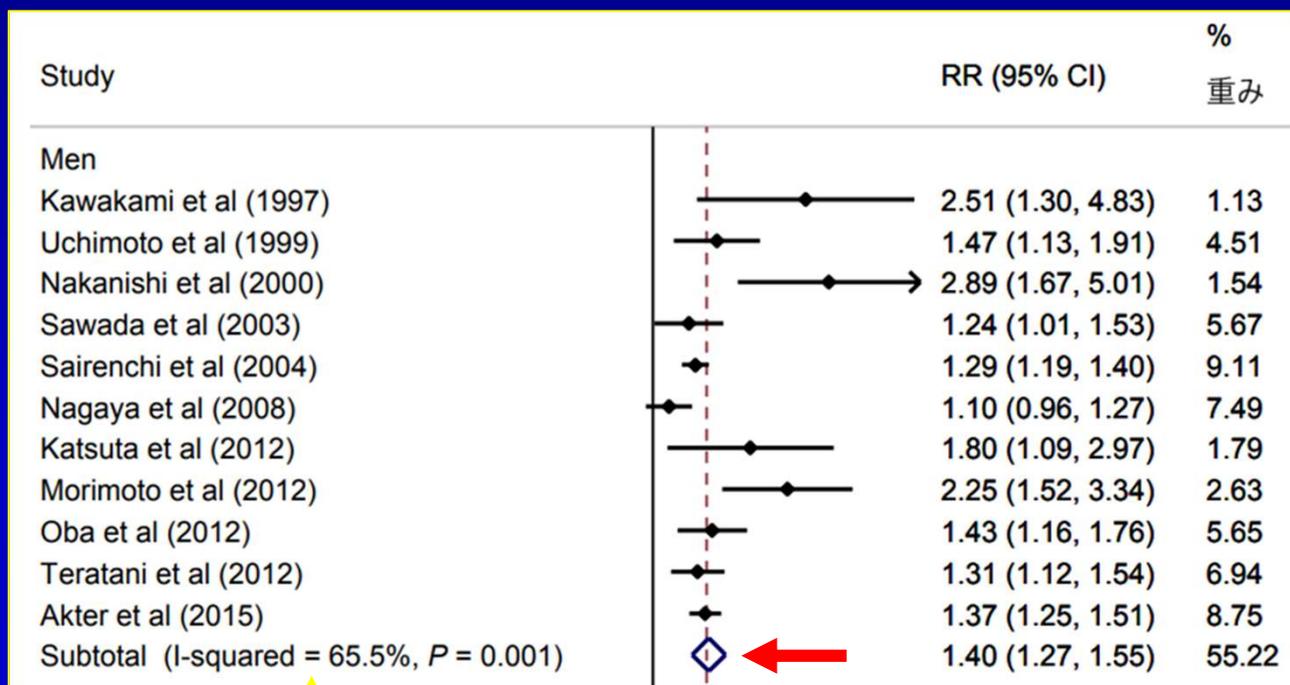
Newcastle-Ottawa Scale (NOS)

Author	Year	Selection				Comparability		Outcome			Total score
		1	2	3	4	5A	5B	6	7	8	
		Exposed cohort truly/somewhat representative	Nonexposed cohort drawn from the same community	Ascertainment of exposure	Outcome of interest not present at start	Cohorts controlled for age and sex	Cohorts controlled for other important factor(s)	Quality of outcome assessment	Follow-up long enough for outcomes to occur	Complete accounting for cohorts	
Kawakami et al	1997		*		*	*	*	*	*		6
Sugimori et al	1998	*	*		*	*	*	*	*		7
Uchimoto et al	1999		*		*	*	*	*	*		6
Nakanishi et al	2000		*		*	*	*	*	*	*	7
Sawada et al	2003		*		*	*	*	*	*		6
Sairenchi et al	2004	*	*		*	*	*	*	*		7
Hayashino et al	2008		*		*	*	*	*	*		6
Nagaya et al	2008		*		*	*	*	*	*		6
Fukui et al	2011		*		*	*	*	*	*		6

観察研究SR/MAの必須事項(結果)

■ 個々の研究と研究全体の点推定値と95%信頼区間等

➤ 全研究の点推定値・信頼区間等を示し、(MAを実施した場合は) 統合値 ← や異質性指標 ← を示しているか？



GRADE

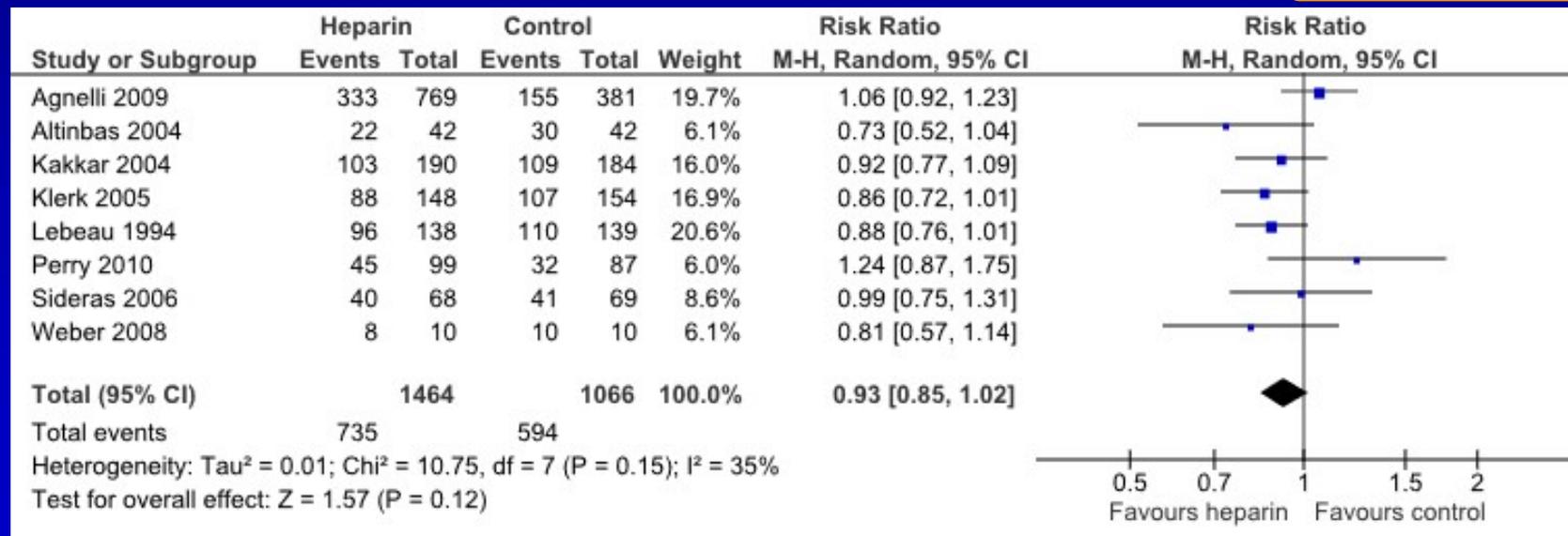
エビデンスの質とは -研究全体の評価-

個々の研究ではなく、
全体で評価

- 系統的レビューにおける定義

効果推定値が正しいという確信の程度

確信の程度は？



Guyatt et al. J Clin Epidemiol. 2011 Apr;64(4):401-6. (No.3)

Guyatt et al. J Clin Epidemiol. 2013 Feb;66(2):151-7. (No.11) ³⁵

GRADE handbook (Updated October 2013)

GRADEでエビデンスの質を吟味する

GRADE

-研究全体の評価-

個々の研究ではなく、
全体で評価

(最初の) エビデンスの 質	グレードを下げる 5ドメイン	グレードを上げる 3ドメイン(観察研究)	(最終的な) エビデンスの質
RCT → 高	<ul style="list-style-type: none"> ・バイアスリスク <ul style="list-style-type: none"> -1 深刻 -2 非常に深刻 ・非一貫性 <ul style="list-style-type: none"> -1 深刻 -2 非常に深刻 ・非直接性 <ul style="list-style-type: none"> -1 深刻 -2 非常に深刻 ・不精確さ <ul style="list-style-type: none"> -1 深刻 -2 非常に深刻 ・出版バイアス <ul style="list-style-type: none"> -1 可能性あり -2 可能性高い 	<ul style="list-style-type: none"> ・効果が大きい <ul style="list-style-type: none"> +1 大きな効果 +2 非常に大きな効果 ・用量反応勾配 <ul style="list-style-type: none"> +1 あり ・全ての交絡を考慮した際 <ul style="list-style-type: none"> +1 効果を減弱させている +1 効果なしの結果であった場合に、効果を増加させる方向に働いている 	<p>高(A) ⊕⊕⊕⊕</p> <p>中(B) ⊕⊕⊕</p> <p>低(C) ⊕⊕</p> <p>非常に低(D) ⊕</p>
観察研究 → 低			

GRADE

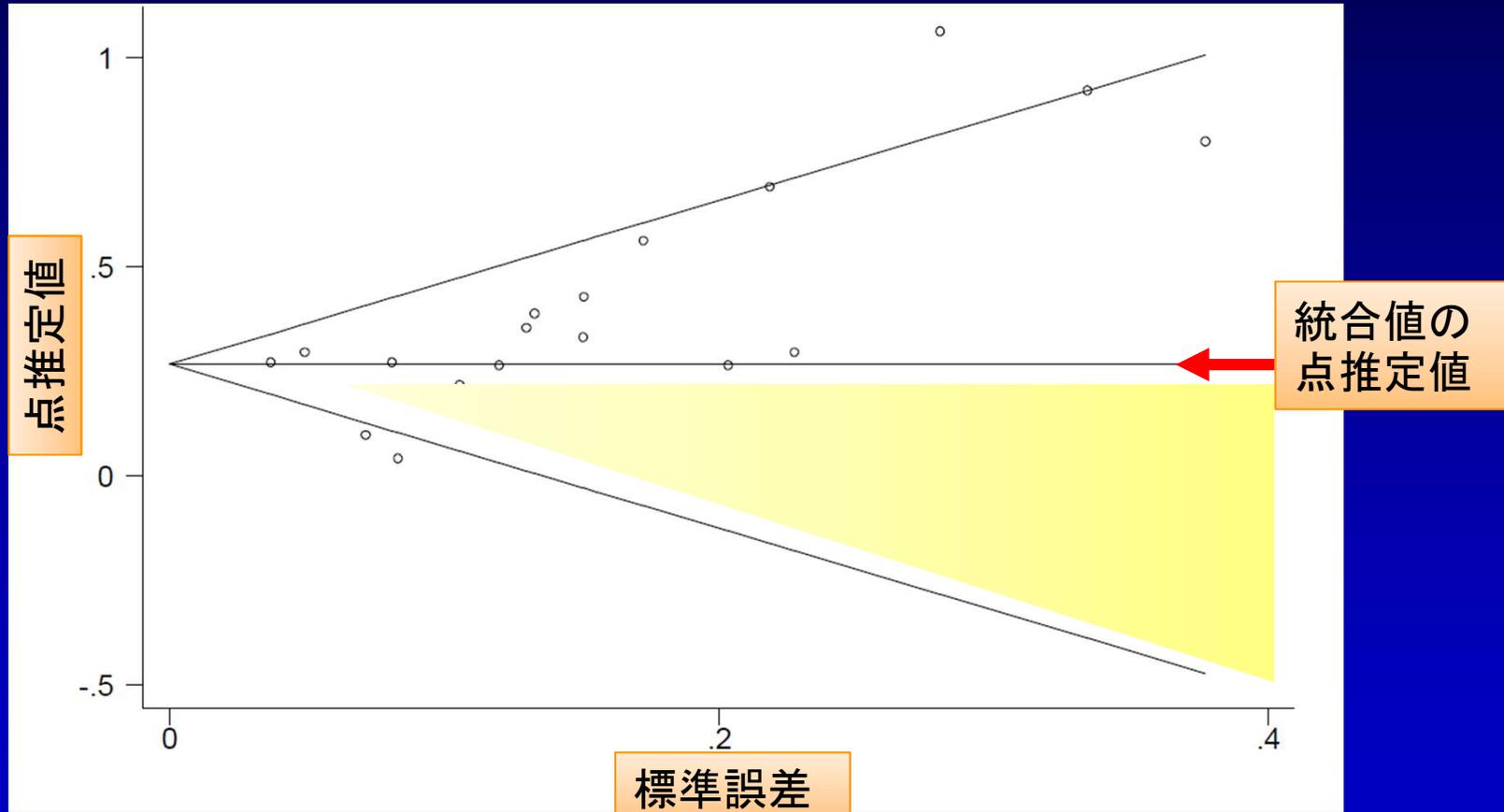
4段階のエビデンスの質

エビデンスの質	効果推定値について
高 (A)	真の効果が効果推定値に近いことに大きな確信がある
中 (B)	効果推定値に対し中等度の確信がある。つまり、真の効果は効果推定値に近いと考えられるが、大きく異なる可能性も否めない。
低 (C)	効果推定値に対する確信性には限界がある。真の効果は効果推定値とは大きく異なるかもしれない。
非常に低 (D)	効果推定値に対し、ほとんど確信が持てない。真の効果は、効果推定値とは大きく異なるものと考えられる

エビデンスの質評価

1. 研究は適切にデザインされ、実施されたか？ **バイアスリスク**
2. 結果は一貫しているか？ **非一貫性**
3. 研究は設定した疑問に直接答えているか？ **非直接性**
4. 結果は精確か？ **不精確さ**
5. 全ての研究を含んでいるか？ **出版バイアス**
 - 観察研究では、用量反応勾配、効果の大きさ、交絡も評価する

出版バイアスの評価 ファンネルプロット



非対称的な分布であると、出版バイアスの存在が疑われる

Take Home Message

- ① SR/MAはPRISMA/MOOSEに従うと書きやすい
- ② プロトコールを作成し、PROSPERO等に登録した方がよい
- ③ 国際誌ではMEDLINE+Embaseがスタンダードである
- ④ 文献検索法を専門家にコンサルトした方がよい
- ⑤ 無理にはメタする必要はない(するべきでない)
- ⑥ 個々の研究の批判的吟味を忘れずに行う
- ⑦ GRADEシステム等で研究全般のエビデンスの質を評価することを必須とする国際誌が増えている

For further readings

1. コクランハンドブック:
http://handbook.cochrane.org/front_page.htm
2. Egger, Smith, Altman “Systematic Reviews in Health Care: Meta-Analysis in Context” (2001).
3. Borenstein et al. “Introduction to Meta-analysis” (2009).
4. 丹後 俊郎 “新版 メタ・アナリシス入門ーエビデンスの統合をめざす統計手法” (2016)
5. Minds診療ガイドライン作成マニュアル 2017 m
第4章:
http://minds4.jcqhc.or.jp/minds/guideline/pdf/manual_all_2017.pdf

SR関連のセミナー

- コクラン日本支部ワークショップ

主催:コクラン日本支部

<http://square.umin.ac.jp/cochranejp/>

- Minds 診療ガイドライン作成ワークショップ

主催:日本医療機能評価機構

<http://minds.jcqhc.or.jp/>