Critical appraisal of systematic review & meta-analysis

Chihaya Koriyama

Department of Epidemiology & Preventive Medicine

Kagoshima University

January 8th, 2021

SUMMARIZING THE EVIDENCE

Clinicians need to establish the context for that one piece of evidence by asking "Have there been other good studies of the same question, what have they shown, and do their results establish a pattern when the studies' scientific strengths and statistical precision are taken into account?"

- Traditional reviews
  - Narrative reviews
- Systematic reviews
- Meta-analysis

Different type of review

Type of Review	Subjective / objective	Reproducibility of literature search	Qualitative/ Quantitative	Statistical analysis
Narrative	Subjective	No	Qualitative	No
Systematic	Objective	Yes	Qualitative	Yes / No
Meta-analysis	Objective	Yes	Quantitative	Yes (necessary)

ELEMENTS OF SYSTEMATIC REVIEW

Elements of a systematic review

1. Define a specific question
2. Find all relevant studies (including unpublished)
3. Select the strongest studies
4. Describe the scientific strength of the selected studies
5. Determine if quality is associated with results
6. Summarize the studies in figures (forest plot) and tables
7. Determine if pooling of studies is justified
8. If so, calculate a summary effect size and confidence interval
9. Identify reasons for heterogeneity if present

Defining a specific question

- Systematic reviews are of specific questions. The elements of specificity have been defined under **PICO(TS)**

  P =patients

  I =intervention

  C =comparison

  O =outcomes

  T =time (follow-up in a cohort study)

  S =study design

An example of critical appraisal worksheet

Center for Evidence-Based Medicine, University of Oxford

What question (PICO) did the systematic review address?

What is best?

The main question being addressed should be clearly stated. The exposure, such as a therapy or diagnostic test, and the outcome(s) of interest will often be expressed in terms of a simple relationship.

Where do I find the information?

The Title, Abstract or final paragraph of the Introduction should clearly state the question. If you still cannot ascertain what the focused question is after reading these sections, search for another paper!

In this paper

Yes	No	Unclear
☐	☐	☐

https://www.cebm.ox.ac.uk/resources/ebm-tools/critical-appraisal-tools

Finding all relevant studies

- **Database**: PubMed, MEDLINE, EMBASE, the Cochrane database of systemic reviews
- Read recent reviews and textbooks
- Seek the **advice of experts**
- Consider articles cited in the articles already found by other approaches
- Information in **non-English language**
- Review **registries of clinical trials** and **funded research** to identify unpublished studies

Limit reviews to scientifically strong, clinically relevant studies

Van den Hoek et al. Statins and the prevention of infections: systematic review and meta-analysis of data from large randomized placebo controlled trials. BJM 2011

632 Potentially relevant → duplicates were excluded

587 → excluded because of review, rationale, study protocol, baseline report, no RCT, no placebo controlled, follow-up<12 months, etc.

38 → excluded because adverse events not specified, data not provided by authors, etc

11 Included in meta-analysis

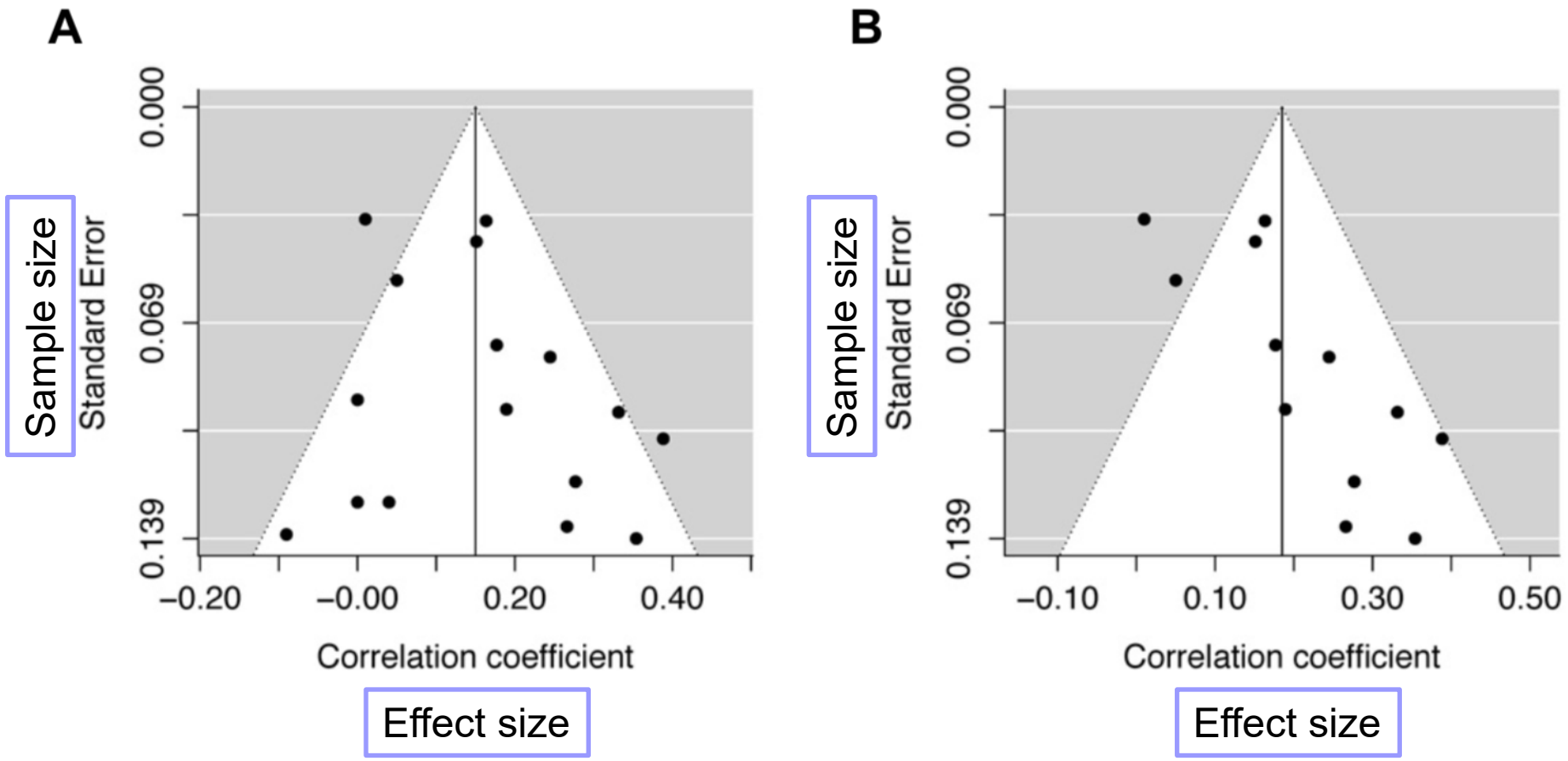
Publication bias

- The articles cited in systematic reviews should include all scientifically strong studies of the question, regardless of publication.
- Publication bias is the tendency for published studies to be systematically different from all completed studies of a question.
- Funnel plots: A scatter plot of the treatment effect estimates from individual trials against a measure of study's precision, usually the standard error (SE).

Funnel plots

Symmetrical

Asymmetrical
→ Possible publication bias

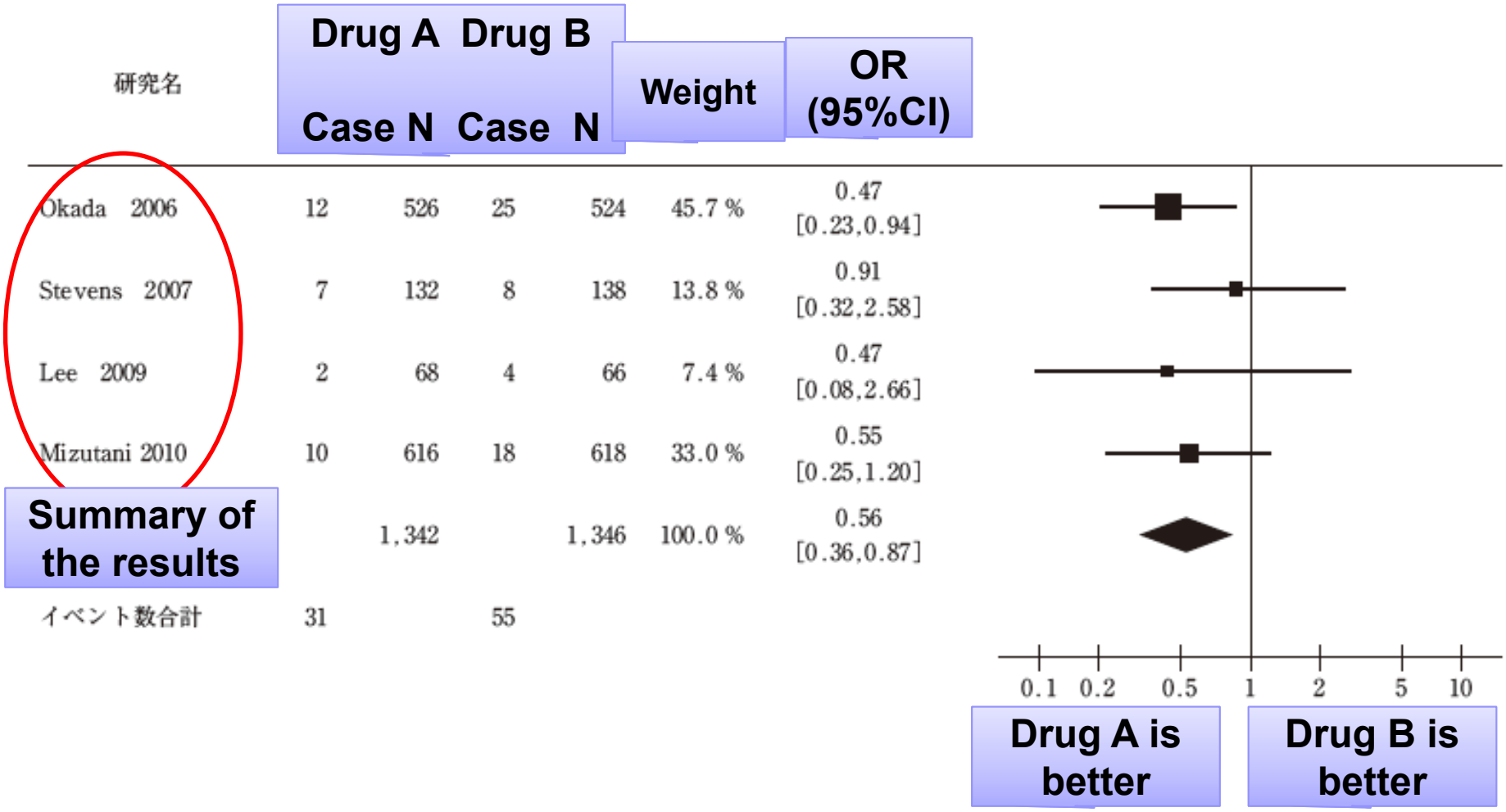


https://towardsdatascience.com/constructing-contour-enhanced-funnel-plots-for-meta-analysis-6434cc8e51d0

Summarizing results

- The results of a systematic review are typically displayed as a forest plot showing the point estimate of effectiveness and confidence interval for each study in the review.
- Point estimates are presented by symbols with their size proportional to the size of the study.

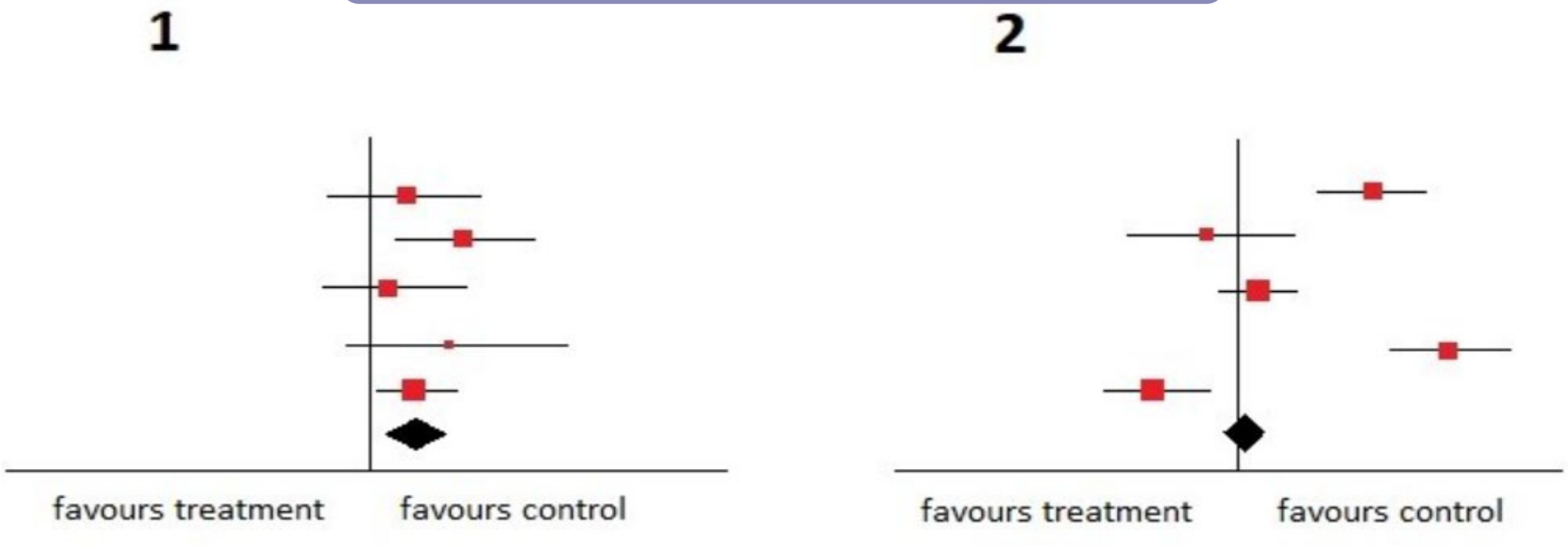
An example of forest plot



Exploring heterogeneity

Heterogeneity can be assessed using the 'eyeball' test or more formally with statistical tests, such as the Cochran Q test.

An example of eyeball test



https://s4be.cochrane.org/blog/2018/11/29/what-is-heterogeneity/

Cochran Q test

H0: the treatments are equally effective
Ha: there is a difference in effectiveness between treatments

If Cochran Q is statistically significant

→There is definite heterogeneity.

If Cochran Q is NOT statistically significant but the ratio of Cochran Q and the degrees of freedom (Q/df) is > 1

→There is possible heterogeneity.

If Cochran Q is NOT statistically significant and Q/df is < 1

→ Then heterogeneity is very unlikely.

How to deal with heterogeneity

- Check your data for mistakes – Go back and see if you maybe typed in something wrong
- Don't do a meta-analysis if heterogeneity is too high – Not every systematic review needs a meta-analysis
- Explore heterogeneity – This can be done by subgroup analysis or meta-regression
- Changing the effect measures – Let's say you use the risk difference and have high heterogeneity, then try out risk ratio or OR

https://s4be.cochrane.org/blog/2018/11/29/what-is-heterogeneity/

STRENGTHS & WEAKNESSES OF META-ANALYSIS

Strengths of meta-analyses

- Meta-analyses, when justified by relatively homogeneous results of component studies, can make many contributions to systematic reviews.
- It can establish that an effect is present or absent with more authority than individual trials.
- Pooling make it possible to estimate effects sizes more precisely so that clinicians can have a better understanding of how big/small the true effect might be.

Weaknesses of meta-analyses

- The temptation to pool quite dissimilar studies, providing a misleading estimate of effects and directing attention away from why differences in effects exist.
- They do NOT include the information based on the biology of disease, clinical experience, and the practical application of best evidence to patient care.

"Which comes closer to the truth, the best individual research or meta-analyses?"

Meta-analyses cannot be better than the scientific strength of the individual studies that they summarize.

INTERNATIONAL GUIDELINES FOR SCIENTIFIC REPORTS

The guidelines for reporting research study

- CONSORT (Consolidated Standards of Reporting Trials)→ RCTs
- STARD (Standards for Reporting of Diagnostic Test Accuracy)→ diagnostic tests
- STROBE (Strengthening the Reporting of Observational Studies in Epidemiology)→ observational studies
- TREND (Transparent Reporting of Evaluations with Nonrandomized Design)→ Non-randomized studies of educational, behavioral, and public health interventions

The guidelines for reporting research study

- QUOROM (Quality of Reporting of Meta-analyses)→ Meta-analyses of RCTs
- MOOSE (Meta-analyses of Observational Studies in Epidemiology)→ Meta-analyses of observational studies
- QUADAS (Quality Assessment of Diagnostic Accuracy Studies)→ Systematic reviews of diagnostic accuracy studies
- GRIPS (Genetic Risk Prediction Studies)→ Genetic risk prediction studies

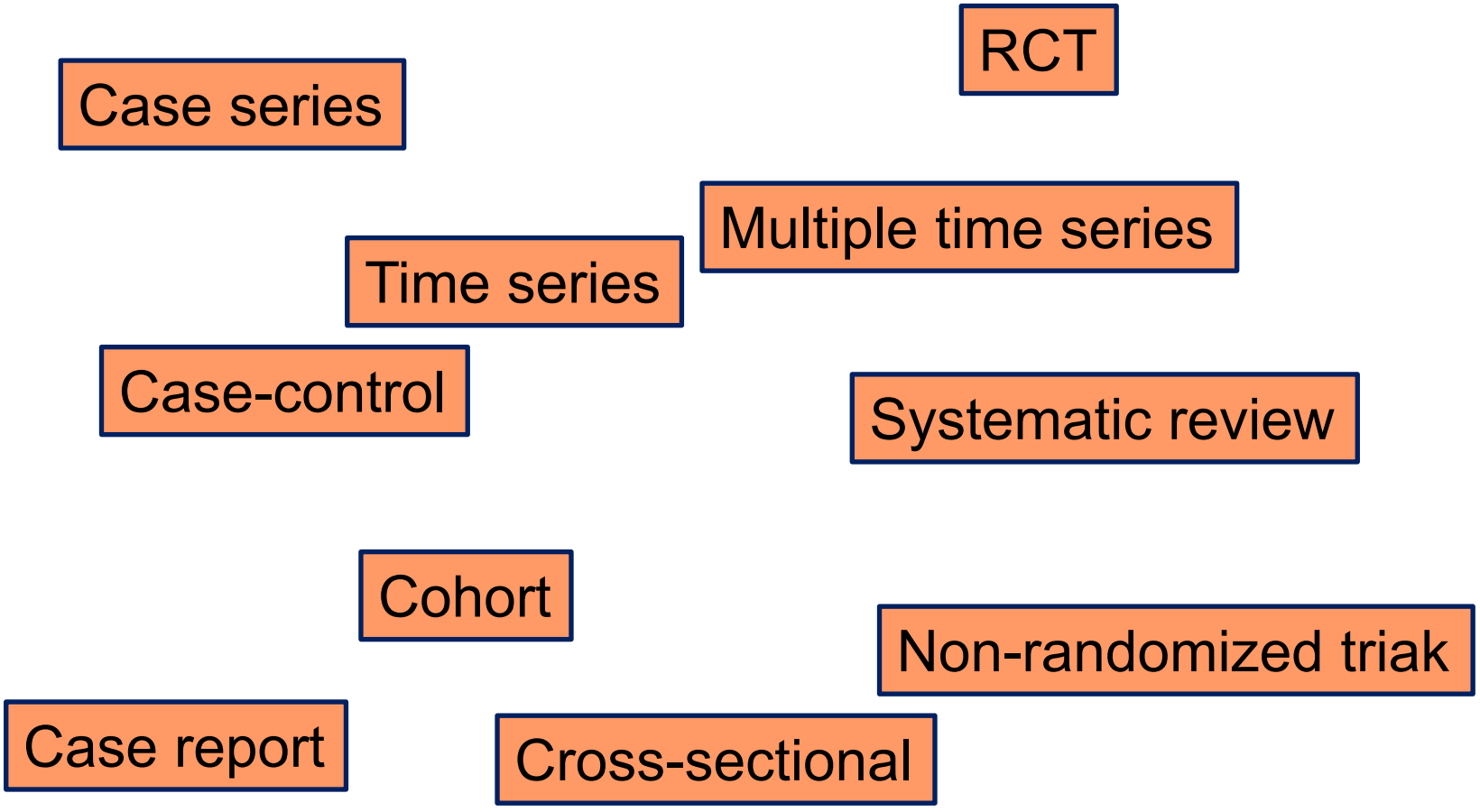
FYI: other worksheets for critical appraisal

- Systematic Reviews
- Diagnostics
- Prognosis
- Randomised Controlled Trials
- Critical Appraisal of Qualitative Studies

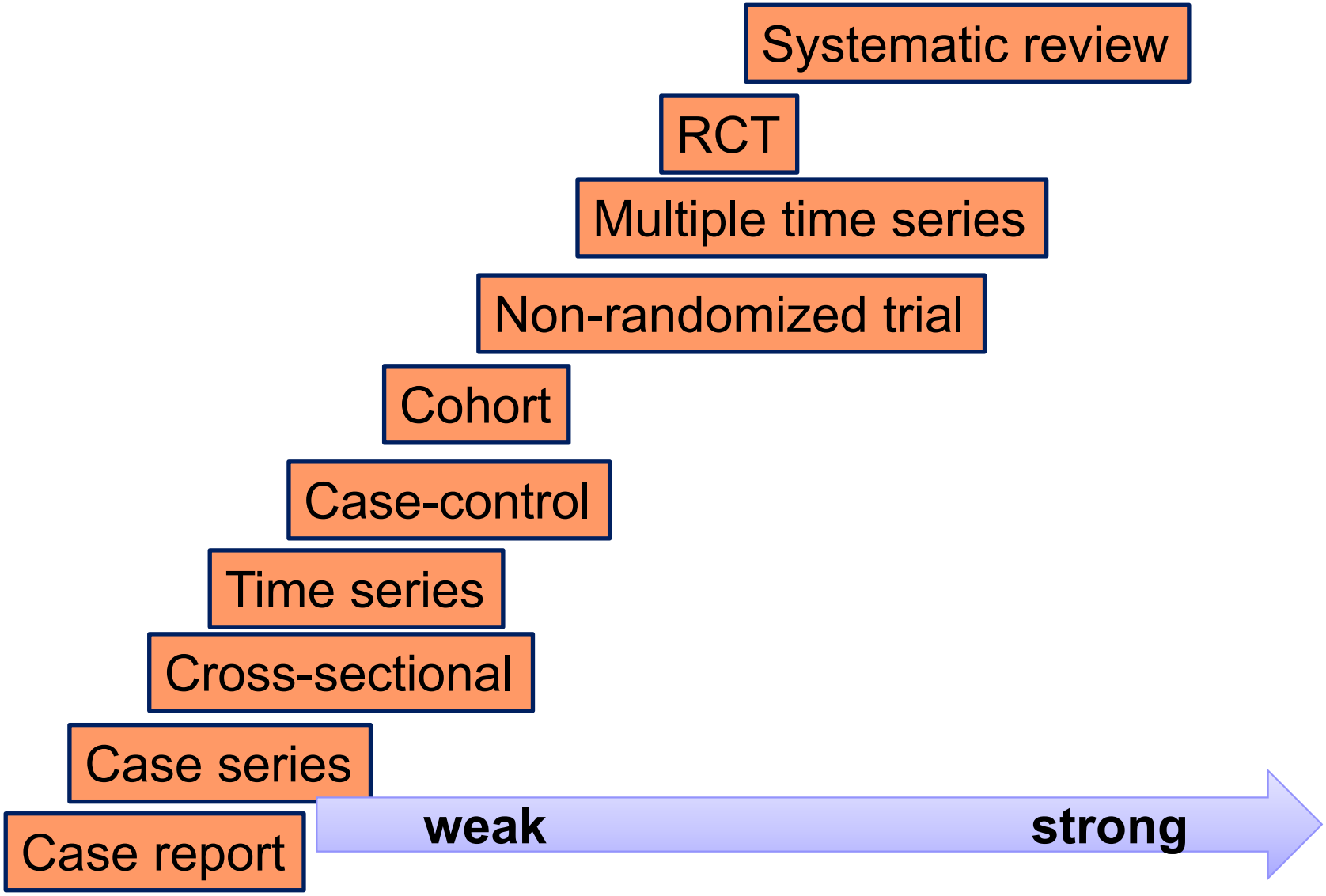
https://www.cebm.ox.ac.uk/resources/ebm-tools/critical-appraisal-tools

STRENGTH OF EVIDENCE BY STUDY DESIGN

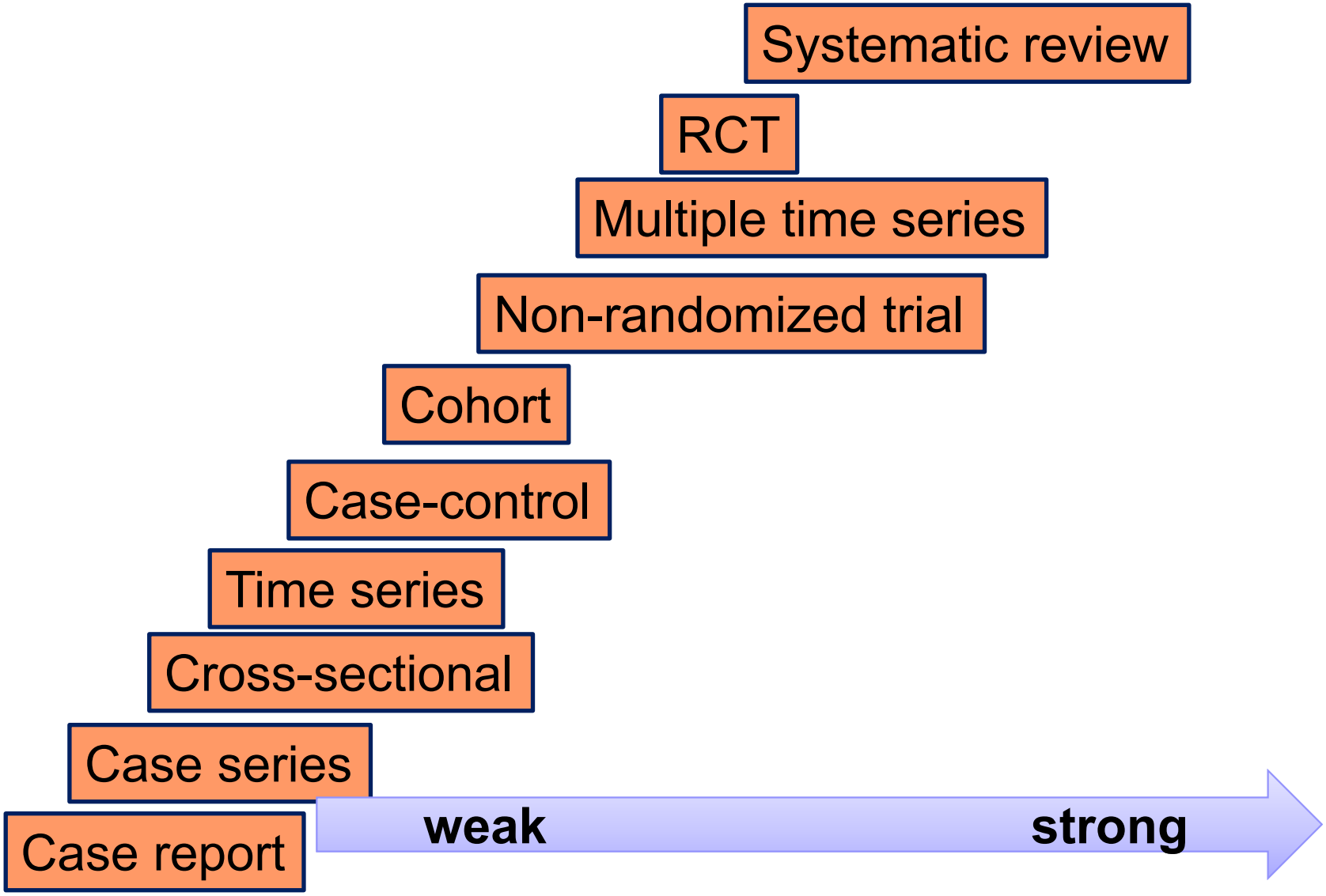
Strength of evidence by study design



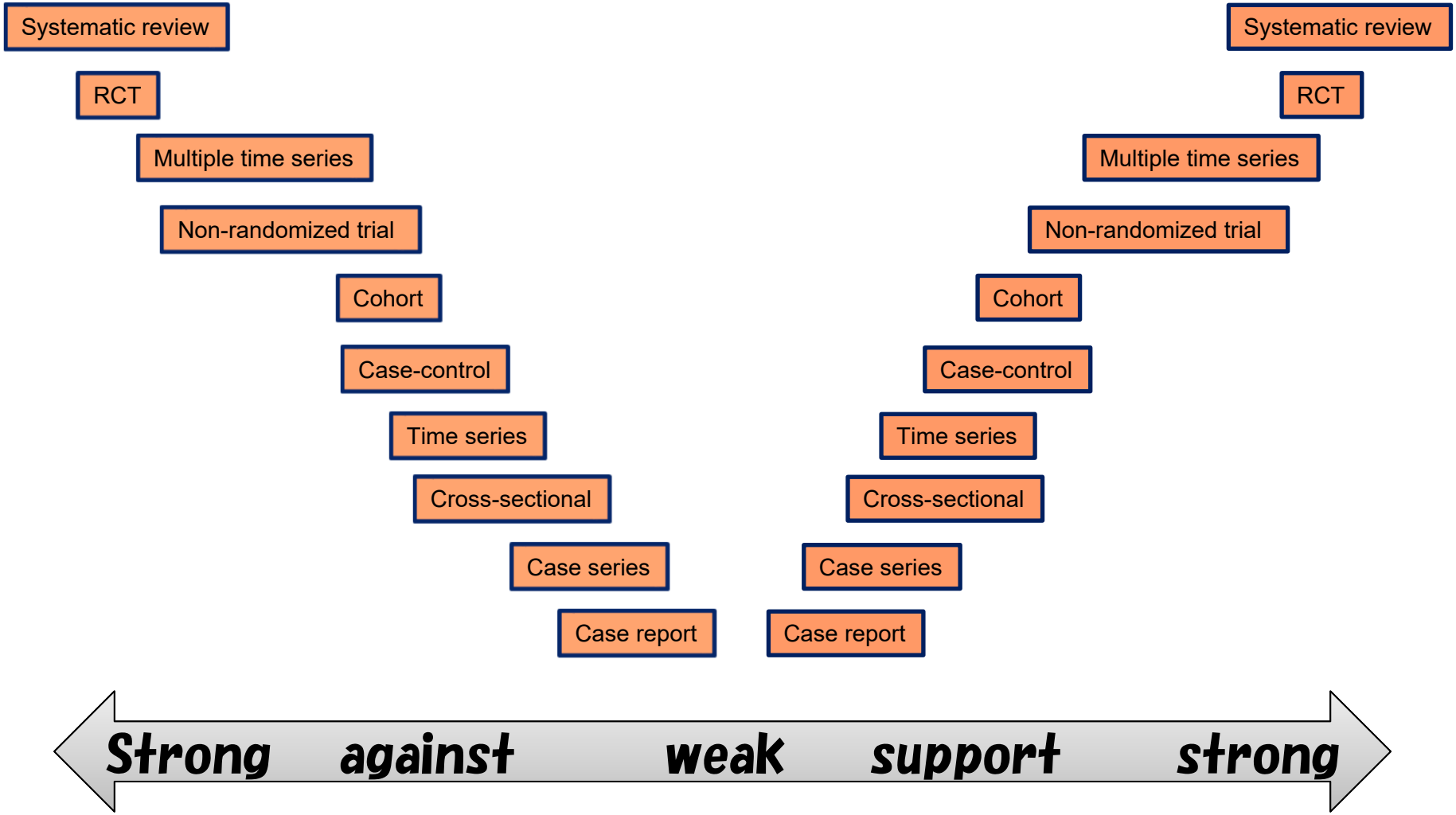
Strength of evidence for causal effect by study design



Strength of evidence **against** causal effect by study design

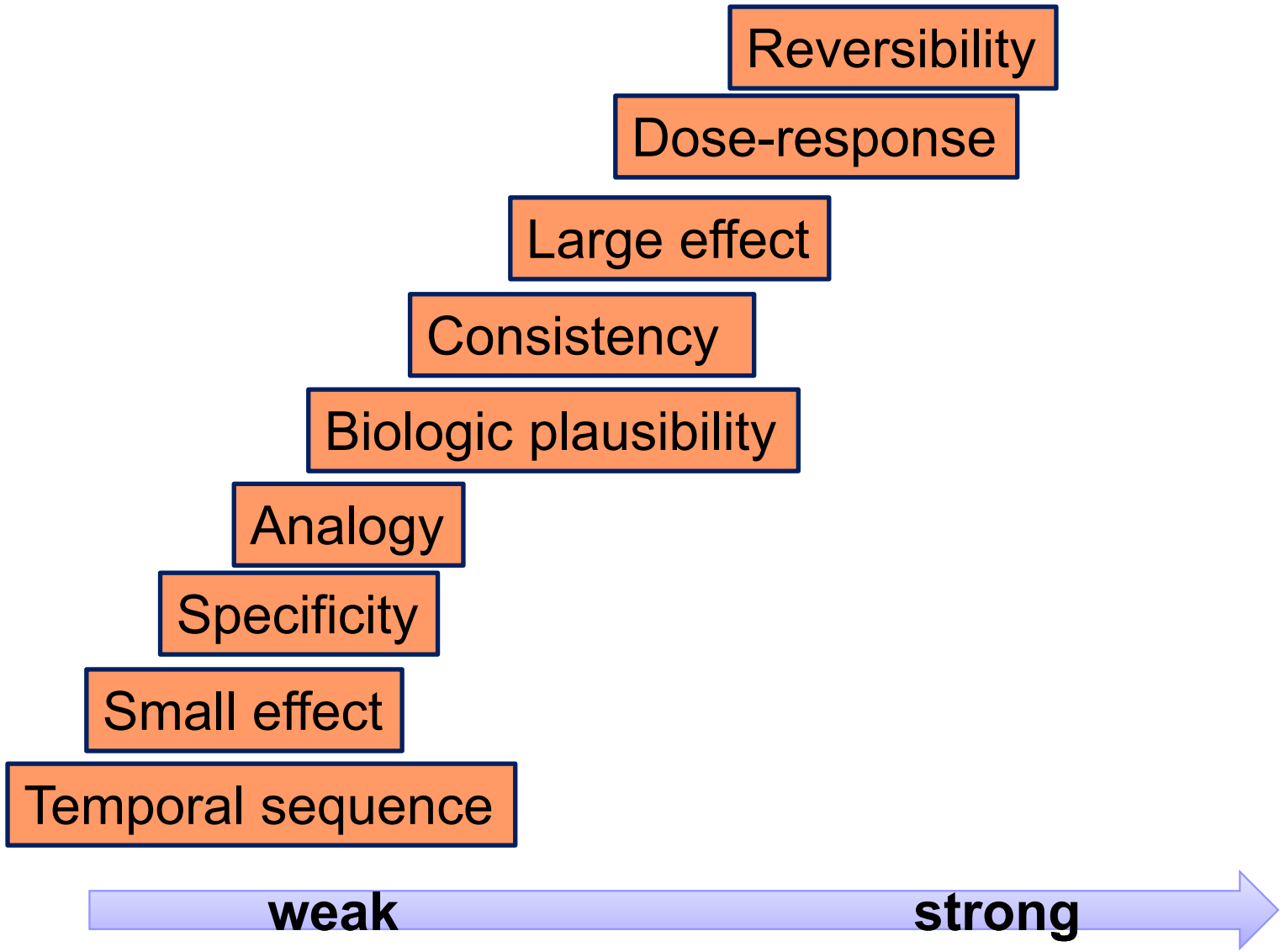


The strength of evidence for a causal relationship by study design is a mirror image of that against



STRENGTH OF EVIDENCE BY FINDINGS

Strength of evidence for causal effect by finding



Criteria for causation

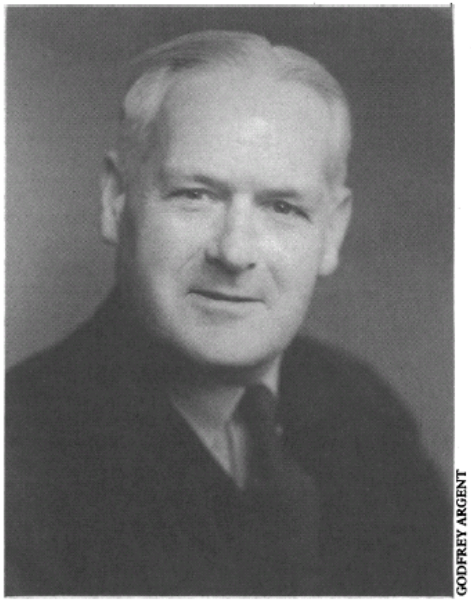
Criteria for the causation between smoking and lung cancer by Surgeon General (1964)

1, strength
2, specificity
3, temporality
4, consistency
5, coherence

Hill's criteria

Sir Hill, a British epidemiologist, added the following criteria to those in Surgeon General Report.

6, dose–response relationship
7, plausibility
8, experimental evidence
9, analogy



Sir Austin Bradford Hill, described as the greatest medical statistician of the twentieth century. He held no degree in either medicine or statistics

With findings, evidence for a causal effect does NOT mirror evidence against an effect