RACE 618: Systematic Review and Meta-analysis

Ammarin Thakkinstian, Ph.D.
Section for Clinical Epidemiology and Biostatistics

e-mail: ammarin.tha@mahidol.ac.th
www.ceb-rama.org
Course learning outcomes

• To know
  • Principle and rational of review methods.

• Be able to
  • Develop a review proposal for systematic review and meta-analysis
  • Perform and conduct a systematic review and meta-analysis
  • Analyse data using basic and advance meta-analyses where appropriate
  • Interpret and report results of systematic review and meta-analysis properly
  • Disseminate results of systematic review to public
Course outline

• Review methodology
• Registration of review protocol
• Grading evidences
Course outline

• Pooling methods
  • Basic meta-analysis
    • Dichotomous outcome
    • Continuous outcome
    • Pooling prevalence/mean
  • Advance meta-analysis
    • Meta-analysis for genetic association studies
    • Network meta-analysis
    • IPD meta-analysis
2 credits

• 1\textsuperscript{st} May – 30\textsuperscript{th} August

• Evaluations
  • Seven assignments (80%)
    • Assignment I: Review topic, rationale, and feasibility*
    • Assignment II: Locate and select studies*
    • Assignment III: Design data extraction & risk of bias
    • Assignment IV: Statistical analysis plan
    • Assignment V: Review proposal
    • Assignment VI: Registration of review protocol at PROSPERO
    • Assignment VII: Present review results/writing manuscript
  • Nine presentations (20%)
    • Presentations
    • Class participations

*Two rounds of presentation
Assignments

• For each assignment
  • Writing assignment
  • Presentations
  • submit assignment 3 days before and 7 days after presentations unless specified

• Final presentation 30\textsuperscript{th} August
• Submit the manuscript to \textbf{Journal by October to December}
What is a systematic review

• A review that has been conducted using a systematic approach in order to minimise biases and random error
Why do we need a systematic review

• Tool for
  • health care practitioners,
  • researchers,
  • policy makers,
  • consumers

who want to keep up with the evidences that are accumulated in their area of interests
Rationale

• More objective appraisal of the evidence than traditional narrative reviews

---

**Narrative review**

• Subjective selection of studies
• Limitation of single or few studies
• Selection bias
• Unhelpful descriptions, e.g., no clear evidence
• A weak relationship, a strong relationship.

**Systematic review**

• Objective selection
• Include identified studies as many as possible, less bias
• More transparent appraisal of evidence
• Allow reader to replicate
• Quantitative conclusion
Rationale

• **Meta-analysis:**
  • Estimates treatment effects
  • Leading to reduces probability of false negative results (increase power of test)
  • Potentially to a more timely introduction of effective treatments.
Rationale

• Exploratory analyses:
  • Subgroups of patients who are likely to respond particularly well to a treatment (or the reverse)

• Systematic review may demonstrate
  • A lack of adequate evidence
  • A gap of knowledge
  • Thus, identify the area where further studies are needed
Terminology

• Systematic review
• Overview
• Meta-analysis
• Research synthesis
• Summarizing
• Pooling
Review Methodology

• Good Rationale
• Clearly state research question
• Objective
• Identify relevant studies/locate studies
• Explicitly describe inclusion & exclusion criteria of studies
• Data extraction
• Statistical analysis
• Results
• Discussion
Review proposal

- Introduction & background & rationale
- Research question/objective
- Review methods
  - Locating studies
  - Selecting studies
    - Inclusion/exclusion criteria
  - Data extraction forms and process
  - Risk of bias assessment
  - Statistical analysis plan
    - Dummy tables/figures
  - Time frame
  - Budget
Rationale

• Why do we need to perform the review
• How were results of previous individual and review studies (if any)
  • Positive results
  • Negative results
• Methodologic issues
  • Sample size/Power of test
  • Previous reviews
    • Narrative reviews?
      • Selective bias
      • Publication bias
      • Pooling effect sizes?
Rationale

• Previous systematic review/s with meta-analysis
  • Methods
    • Selection bias?
    • Pooling appropriately?
    • Number of studies?
    • Number of relevant outcomes?
    • Number of treatments?
  • Number of publications since previous published?
Management of Chronic Prostatitis/Chronic Pelvic Pain Syndrome
A Systematic Review and Network Meta-analysis

Magnitude of problem

• Prostatitis is a common condition, with an estimated prevalence in the community of about 9%, and accounts for nearly 2 million ambulatory care encounters annually in the United States.

• Symptoms of CP/CPPS can diminish quality of life and impair physical and psychological function.
• The etiology of CP/CPPS is uncertain but may include inflammatory or noninflammatory etiologies.\textsuperscript{6,7,8}

• An inciting agent may cause inflammation or neurological damage in or around the prostate and lead to pelvic floor neuromuscular and/or neuropathic pain.

• Predisposing factors for CP/CPPS may include heredity, infection, voiding abnormalities, hormone imbalance, intraprostatic reflux, immunological or allergic triggers, or psychological traits.
A wide variety of therapies including α-blockers, antibiotics, anti-inflammatory medications, and other agents (eg, finasteride, phytotherapy, and gabapentinoids) are routinely used.

**Rationale**

- However, the efficacy of these treatments is controversial, partly because many clinical trials testing these therapies have been small, with little statistical power to detect treatment effects.
- To date, only 1 systematic review and 1 meta-analysis of α-blockers vs placebo of which we are aware have been performed for treatment of CP/CPPS.
• We therefore performed a systematic review and network meta-analysis mapping all treatment regimens, with 2 aims.
  • To compare total symptom, pain, voiding, and quality-of-life scores at the end of therapy with α-blockers (the most commonly evaluated therapy for CP/CPPS), other active drugs, or placebo.
  • To compare rates of responses to therapies available for treating CP/CPPS.
Reno-protective effects of renin-angiotensin system blockade in type 2 diabetic patients: a systematic review and network meta-analysis

Background & rationale

• Diabetic nephropathy is a significant health and economic burden across the world.

• The prevalence of micro- and macro-albuminuria in type 2 diabetes is as high as 37–40% in western countries and 57.4–59.8% in Asian countries

• Preventive treatments have been prescribed for T2D with or without hypertension with the aim of lowering BP, and delaying or even preventing the progression of diabetic nephropathy
• The reno-protective effects of renin–angiotensin system (RAS) blockers in type 2 diabetes have been controversial \([8]\).

• A few systematic reviews have been conducted \([9–11]\), but
  
  • these reviews pooled studies with mixed populations of participants with type 1 and type 2 diabetes, with and without diabetic nephropathy, and
  
  • the focus was mainly on surrogate rather than clinical outcomes
Rationale (cont.)

• We therefore conducted a systematic review and meta-analysis with aims at
  • Comparing the effects of ACE inhibitor (ACEI)/angiotensin II receptor blocker (ARB) with other antihypertensive drugs and placebo
  • On outcomes of
    • ESRD,
    • Doubling of serum creatinine,
    • Microvascular complications,
    • Micro- and macroalbuminuria and
    • Regression of albuminuria
Risk prediction models of breast cancer: a systematic review of model performances

Anothaisintawee T, Teerawattananon Y, Wiratkapun C, Kasamesup V, Thakkinstian A.
• Breast cancer is the most common cancer found in women across the world, accounting for 23% of all new cancers in women [1].

• Although around two-thirds of new cases are in developed countries, it is the most common cancer in women in developing countries.

• An early detection screening using mammography can improve the survival rate of patients using strong evidence suggested from a meta-analysis [3].

• The organized breast cancer screening programs using mammography have therefore been well established in developed countries in Europe and North America.
In contrast, the majority of women living in developing countries have never had access to mammography because of:

- the severe shortage of human resources and
- infrastructures that have been needed for a few decades to fill the gap.

Screening to every woman is therefore not feasible in most developing countries.

But identifying target women with relatively higher risk of developing breast cancer looks to be a promising alternative.
The number of risk prediction models has been increasingly developed, for estimating about breast cancer in individual women. However, those model performances are questionable, i.e., poor to fair performances. We therefore have conducted a study with the aim to systematically review previous risk prediction models:

- How the models were constructed?
- How many variables and what were they?
- Study design?
• Were those models validated?
  • Internal and external validations?
• Model’s performance
  • Calibration
  • Discrimination

• The results from this review help to identify the most reliable model and indicate the strengths and weaknesses of each model for guiding future model development.
Good research question

• Evidence-base Medicine (EBM)
  • Patient/Population
  • Intervention/Exposure
  • Comparator
  • Outcome
  • PICO
Flow diagram of review for conducting further study
Research question

Treatments

• CP/CPPS
  • Is alpha-blocker is better in improving total symptom, pain, voiding, and quality of life than antibiotics in CP/CPPS patients?
  • Among active treatments, which treatment regimens are better in improving symptoms in CP/CPPS patients?
Research question

• Preeclampsia
  • Does calcium or vitamin D supplementation decrease risk of preeclampsia occurrence in pregnancy when compares to placebo?
  • Among calcium, vitamin D, and calcium plus vitamin D supplementations, which one is better in prevention of preeclampsia in pregnancy?
Research question

• Diagnostic studies
  • How are performances of Berlin and Stop-Bang questionnaires comparing with the standard polysomnography in screening obstructive sleep apnea in pregnancy

• Observational studies
  • Does sleep duration associate with type two diabetes and its progression in general adults?
  • Is there association between VDR and BMD/osteoporosis in women?
Research question

• Prediction model of breast cancer risk
  • Is there any prediction model of breast cancer occurrence
  • If yes, what are they?, how had they been developed
    • Setting
    • Time
    • Predictors and measures
    • Type of model
  • How were their model’s performances in internal and external data?
Locate studies

1. Define sources of database
   • MEDLINE
     - 1949 to present
     - Over 16 million references
     - Completed references are added each day from Tuesday through Saturday
     - Cover 5200 worldwide journals in 40 languages
     - Uses medical subject heading (MeSH) for index
     - Includes biomedicine and health science journals
       - English abstracts for 79% on references
       - 90% are English language articles
       - 47% of journals covered are published in the US
     - PubMed available free of charge

Define sources of database

EMBASE

- Over 12 million records from 1974-present
- More than 600,000 records added annually
- Covers over 4,800 active peer-reviewed journals published in > 70 countries/ 30 languages
- uses EMTREE for indexing
- includes English abstracts for 80% of references
- daily update, within two weeks of receipt of the original journal
- Produced by Elsevier, no free version available
Define sources of database

Scopus (launched in November 2004)

- 18,000 titles
  - 16,500 peer-reviewed journals (1,200 Open Access journals)
  - 600 trade publications
  - 350 book series
  - 3.6 million conference papers (~10%) from proceedings and journals
    - Medical Science ~2.9%
    - Biological Science ~ 2.7%
    - Chemical Science ~ 1.9%
• 41 million records
  • 21 million records with references back to 1996
  • 20 million records 1823-1996

• 318 million scientific web pages

• 23 million patent records from five patent offices
  • World Intellectual Property Organization (WIPO)
  • European Patent Office
  • US Patent Office
  • Japanese Patent Office
  • UK Intellectual Property Office
• “Articles-in-Press” from over 3,000 journals
  • Cambridge University Press
  • Elsevier
  • Springer / Kluwer
  • Karger Medical and Scientific Publishers
  • Nature Publishing Group (NPG)
  • The Institute of Electrical and Electronics Engineers (IEEE)
  • BioMed Central (BMC)
  • Lippincott, Williams & Wilkins (LWW)
Coverage by region

Number of Scopus titles by geographical region (October 2009)

Percentage of journals in Scopus based on geographical regions (January 2010)
Coverage across subject areas

- **Health Sciences** (100% Medline)
  - Nursing
  - Dentistry
  - ...

- **Physical Sciences**
  - Chemistry
  - Physics
  - Engineering
  - ...

- **Life Sciences**
  - Neuroscience
  - Pharmacology
  - Biology
  - ...

- **Social Sciences**
  - Psychology
  - Economics
  - Business
  - Arts & Humanities
  - ...

Number of journal titles by broad subject area.
Note: Journal titles may belong to more than one subject area.
Define sources of database

- The Cochrane Controlled Trials Register (CCTR)
- ClinicalTrials.gov
- HUGE NET Review
- Reference lists
- Personal communication with expert in the field
Define source of database

• Gray literatures
  • Information that falls outside the mainstream of published journal and monograph literature, not controlled by commercial publishers

• Sources from NSH library*:

• WorldCat - 1.5 billion items in this collection of library catalogs
• Google Scholar - Search scholarly literature across many disciplines and sources, including theses, books, abstracts and articles.

*PS accessed 22/04/2018
Gray literatures

- **Gray Source Index**
- **AHRQ** - agency for healthcare research and quality
- **World Health Organization** - providing leadership on global health matters, shaping the health research agenda, setting norms and standards, articulating evidence-based policy options, providing technical support to countries and monitoring and assessing health trends.
- **List Gray Literature Producing Organizations** - from the New York Academy of Medicine, includes government and private sector
Locate studies

2. Define the software & version used for searching
   - PubMed
   - Ovid
   - Scopus
   - EmBase
   - Web of Science
National Center for Biotechnology Information


A novel single-nucleotide polymorphism of the Fcgamma receptor IIa gene is associated with genetic susceptibility to systemic
Search Results

1. Role of LINGO1 polymorphisms in Parkinson's disease
   Haubenberger D., Hotzy C., Pirker W., Katzenschläger R., Brücke T., Zimprich F., Auff E., Zimprich A.
   *Movement Disorders* 2009 24:16 (2404-2407)
   Embase, MEDLINE, Abstract, Index Terms

2. Relationship between glucose-6-phosphate dehydrogenase gene mutations and neonatal jaundice in
   Nanjing, Guangxi
   *Chinese Journal of Contemporary Pediatrics* 2009 11.12 (970-972)
3. Defines searching terms

- Combinations of search terms based on PICO
  - Patient
  - Intervention: treatment/study factor
  - Comparator
  - Outcome of interest
- Specify period of searching
- Plan for update searching
Example

- VDR & BMD/Osteoporosis (J Bone Miner Res. 2004;19(3):419-28.) Intervention/exposure
- P
  - Women
  - Females
- I/E
  - Vitamin D receptor
  - VDR
  - Genotype
  - Allele
  - Polymorphism
  - Locus
Outcome

- Bone mineral density
- BMD
- Bone density
- Osteoporosis
- Fracture
Example

- **VDR** BMD/Osteoporosis(J Bone Miner Res. 2004;19(3):419-28.)

1. vitamin D receptor or VDR (MeSH)
2. genotype(s) or allele(s) or polymorphism(s) (MeSH)
3. bone mineral density or BMD or bone density (MeSH)
4. low bone mineral density or low density (textword)
5. osteoporosis (MeSH)
6. fracture (MeSH)
7. 1 and 2 and 3
8. 1 and 2 and 4
9. 1 and 2 and 5
10. 1 and 2 and 6
PubMed Advanced Search Builder

Use the builder below to create your search

**Builder**

- **All Fields**
- **AND**

**History**

There is no recent history
Selecting studies

• Clearly define inclusion & exclusion criteria

• Inclusion criteria base on PICO
  • P: Type of patients
    • Children, adults
    • Specific type of disease
      • T2D, CKD, CP/CPPS IIIA
  • I: Treatment or exposure or gene or prediction model
  • C: Comparator (if needed)
  • Outcome
General criteria

• Study design
  • randomized controlled trial
  • observational studies (cohort, case-control, cross-sectional studies)

• Full paper Languages
  • English, French, others

• Period of publications

• Multiple publications of the same studies, choose the recent one or the one has provided more completeness of data
Exclusion

- Incompleteness of information
  - Contact authors at least two times for incomplete data

Design coding for ineligibility criteria

- Not studied patients of interest
- Not the outcome of interests
- Not the intervention of interests
- Study design
  - Not comparative studies, no control group
  - Not RCTs
- Review studies
  - Narrative review, systematic review
Selecting studies

- Merge studies identified from databases using reference manager (e.g. Endnote)
  - Remove duplicates

- Two reviewers independently select studies
  - Screen title/abstract to remove non-relevant studies based on eligibility criteria
  - Access and review full papers
  - Computerize review results
• Check reference lists if required
• Contact author if needed
• Final decision

• Perform update searching every 1-3 months while doing a review
  • Construct auto-search once search strategies are finalised for each search engine
  • Perform update selection of study flow if there is at least one study published since last search
Figure 1. Flow of information through the different phases of a systematic review.
doi:10.1371/journal.pmed.1000097.g001
Study selection example

- Participants with CP/CPPS categories IIIA or IIIB
- Any pair of the following interventions:
  - $\alpha$-blockers,
  - antibiotics,
  - steroidal and nonsteroidal anti-inflammatory drugs,
  - finasteride, glycosaminoglycans, phytotherapy, gabapentinoids, and placebo.
- Any of the following outcomes:
  - pain scores, voiding scores, quality-of-life scores, and total symptom scores.
• The full article could be retrieved
• Had sufficient data for extraction, including number of patients, means and standard deviations of continuous outcomes in each group, and/or numbers of patients per group for dichotomous outcomes.
• For trials with multiple publications, we selected the publication with the most complete information.
• Disagreements in selection were resolved by discussion and consensus.
Anothaisintawee, T. et al. JAMA 2011;305:78-86
Data extraction (DE)

- At least two reviewers
- Design DE forms (DEF), pilot, & revise the DEFs
- General characteristic of article
  - Study ID,
  - First Author’s & corresponder’s surnames & their emails
  - Year & source of publication
- The study characteristics
  - Setting
  - Country
  - Study design (RCT, CS, CC, CrS)
  - Type of study phase (for prediction model)
  - Type of studied subjects
    - Ethnicity, setting
    - Children, adults, pregnancy
    - Postmenopause, premenopause
    - General vs specific disease
• **Patients**
  • Demographic and clinical features of studied participants that might associate with outcomes
    • mean age, gender, BMI, smoking, underlying diseases
• **Methods/criteria used for measurement**
  • Outcome
  • Studied factor
• **Interventions/exposure/test**
  • Treatments
    • Dosage/day
    • Period/course of treatments
    • Route
  • Scanners
    • Version
• **Lab tests**
• **Questionnaire & cutoff**
Data for pooling

• Dichotomous outcome
  • Frequency data
  • Summary statistic data
    • Odds ratio (OR) & 95% CI
    • Risk ratio (RR) & 95% CI

• Continuous outcome
  • Summary data
    • Mean and standard deviation
  • Summary statistic data
    • Beta coefficient, i.e., mean difference & 95% CI

• Time to event data
  • Numbers of event & person-time at risk
  • Summary statistic data
    • Hazard ratio & 95% CI
    • Kaplan-Meier curve
Data for pooling for dichotomous outcome

- Contingency table of studied factors/interventions versus outcomes (rxc)

<table>
<thead>
<tr>
<th>Group</th>
<th>Disease</th>
<th>n</th>
<th>Incidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rx (Exp+)</td>
<td>Yes</td>
<td>a</td>
<td>n₁</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>b</td>
<td>a/n₁</td>
</tr>
<tr>
<td>Placebo (Exp-)</td>
<td>Yes</td>
<td>c</td>
<td>n₂</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>d</td>
<td>c/n₂</td>
</tr>
</tbody>
</table>
Data for pooling for dichotomous outcome

- Summary statistic data

<table>
<thead>
<tr>
<th>Intervention/Exposure</th>
<th>ES</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR/RR</td>
<td>LL</td>
</tr>
<tr>
<td>Rx (Exp+)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo (Exp-)</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

- ES, effect size
Data for pooling for continuous outcome

- **Continuous outcome**
  - $n$, mean, SD

<table>
<thead>
<tr>
<th>Group</th>
<th>$n$</th>
<th>mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$n_1$</td>
<td>mean$_1$</td>
<td>SD$_1$</td>
</tr>
<tr>
<td>B</td>
<td>$n_2$</td>
<td>mean$_2$</td>
<td>SD$_2$</td>
</tr>
</tbody>
</table>
• Summary statistic data for continuous outcome

<table>
<thead>
<tr>
<th>Intervention/Exposure</th>
<th>ES</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rx (Exp+) vs Placebo (Exp-)</td>
<td>beta</td>
<td>LL</td>
</tr>
</tbody>
</table>

• PS: Beta, unstandardized beta coefficient
Data for pooling for time to event outcome

<table>
<thead>
<tr>
<th>Intervention/Exposure</th>
<th>Person-time at risk</th>
<th>No. event</th>
<th>Incidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rx (Exp+)</td>
<td>(scale)*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo (Exp-)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Must be the same scale, e.g., year, month, etc.
Data for pooling for time to event outcome

• Summary statistic data

<table>
<thead>
<tr>
<th>Intervention/Exposure</th>
<th>ES</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>HR</td>
</tr>
<tr>
<td>Rx (Exp+)</td>
<td></td>
<td>LL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UL</td>
</tr>
<tr>
<td>Placebo (Exp-)</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

• ES, effect size
Risk of bias in individual studies

- Quality Assessment (QA)
- Consider internal & external validity
Risk of bias (cont.)

- RCT
  - The Cochrane Collaboration’s tool for assessing risk of bias 2009
    - Preferred reports of items for systematic review and meta-analysis-PRISMA guideline
<table>
<thead>
<tr>
<th>Domain</th>
<th>Description</th>
<th>Review authors’ judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence generation.</td>
<td>Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.</td>
<td>Was the allocation sequence adequately generated?</td>
</tr>
<tr>
<td>Allocation concealment.</td>
<td>Describe the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment.</td>
<td>Was allocation adequately concealed?</td>
</tr>
<tr>
<td>Domain</td>
<td>Description</td>
<td>Review authors’ judgement</td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Blinding of participants, personnel and outcome assessors</td>
<td>Assessments should be made for each main outcome (or class of outcomes). Describe all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.</td>
<td>Was knowledge of the allocated intervention adequately prevented during the study?</td>
</tr>
<tr>
<td>Incomplete outcome data</td>
<td>Assessments should be made for each main outcome (or class of outcomes). Describe the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. State whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomized participants), reasons for attrition/exclusions where reported, and any re-inclusions in analyses performed by the review authors.</td>
<td>Were incomplete outcome data adequately addressed?</td>
</tr>
<tr>
<td>Domain</td>
<td>Description</td>
<td>Review authors’ judgement</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Selective outcome reporting.</td>
<td>State how the possibility of selective outcome reporting was examined by the review authors, and what was found.</td>
<td>Are reports of the study free of suggestion of selective outcome reporting?</td>
</tr>
<tr>
<td>Other sources of bias.</td>
<td>State any important concerns about bias not addressed in the other domains in the tool.</td>
<td>Was the study apparently free of other problems that could put it at a high risk of bias?</td>
</tr>
<tr>
<td></td>
<td>If particular questions/entries were pre-specified in the review’s protocol, responses should be provided for each question/entry.</td>
<td>Premature trial termination</td>
</tr>
<tr>
<td></td>
<td>Trial methodology</td>
<td>Post-randomization exclusion</td>
</tr>
<tr>
<td></td>
<td>Statistical analysis</td>
<td>Unbalance baseline characteristics</td>
</tr>
<tr>
<td></td>
<td>Adequately describe methods of data analysis</td>
<td>Adequately describe methods of data analysis</td>
</tr>
<tr>
<td></td>
<td>- use per-protocol analysis, modified ITT</td>
<td>- use per-protocol analysis, modified ITT</td>
</tr>
</tbody>
</table>
Risk of bias assessment

<table>
<thead>
<tr>
<th>Author</th>
<th>Adequate sequence generation</th>
<th>Adequate allocation concealment</th>
<th>Blinding</th>
<th>address incomplete outcome data</th>
<th>Selective outcome report</th>
<th>Free of other bias</th>
<th>Description of other bias</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Risk Of Bias In Non-randomised Studies-of Interventions (ROBINS-I)

• For intervention studies where interventions are not randomly allocated.

• Seven domains are considered
  • Before interventions
    • Confounding
    • Selection of patients into the study
  • At interventions
    • Classification of interventions
ROBINS-I

• After interventions
  • Deviation from intended interventions
  • Missing data
  • Measurements of outcomes
  • Selective outcome report

• The first three domains are totally different from assessments of RCT because randomisation can protect against bias before/at randomization

• The last four domains overlapped with RCT because randomization does not protect bias after randomisation
ROBINS-I

• Response options for each domain
  • Yes, Probably yes
  • No, Probably no
  • No information

• Overall risk of bias judgment
  • Low risk
    • All seven domains are low risk of bias
• Moderate risk
  • The study is judged to be low and moderate risks for all domains

• Serious risk
  • The study is judged to be serious risk of bias at least one of all domains

• Critical risk
  • The study is judged to be critical risk of bias at least one of all domains
Observational studies

• NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE (NOS)
• Risk/association studies
• Cohort studies
  • Selection of cohorts
  • Comparability of cohorts
  • Assessment of outcome
• Items
  • Selection (4)
  • Comparability (1)
  • Exposure (3)

NOS

• Case-Control studies
  • Selection of case and controls
  • Comparability of cases and controls
  • Ascertainment of exposure

• Items
  • Selection (4)
  • Comparability (1)
  • Exposure (3)
Rating

• Grade ‘high’ quality as a ‘star’
• A maximum of one ‘star’ for each item within the ‘Selection’ and ‘Exposure/Outcome’ categories; maximum of two ‘stars’ for ‘Comparability’
• Prognostic studies
  • Quality in prognostic study (QUIPS)
  • Study participants
  • Study attrition
  • Prognostic factor measurement
  • Outcome measurement
  • Study confounding
  • Statistical analysis and report
  • Each domain is graded as low, moderate, and high risk of bias
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal study or characteristics of unbiased study</td>
<td>The study sample adequately represents the population of interest</td>
<td>The study data available (i.e., participants not lost to follow-up) adequately represent the study sample</td>
<td>The PF is measured in a similar way for all participants</td>
<td>The outcome of interest is measured in a similar way for all participants</td>
</tr>
<tr>
<td>Prompting items and considerations†</td>
<td>a. Adequate participation in the study by eligible persons</td>
<td>a. Adequate response rate for study participants</td>
<td>a. A clear definition or description of the PF is provided</td>
<td>a. A clear definition of the outcome is provided</td>
</tr>
<tr>
<td></td>
<td>b. Description of the source population or population of interest</td>
<td>b. Description of attempts to collect information on participants who dropped out</td>
<td>b. Method of PF measurement is adequately valid and reliable</td>
<td>b. Method of outcome measurement used is adequately valid and reliable</td>
</tr>
<tr>
<td></td>
<td>c. Description of the baseline study sample</td>
<td>c. Reasons for loss to follow-up are provided</td>
<td>c. Continuous variables are reported or appropriate cut points are provided</td>
<td>c. The method and setting of outcome measurement is the same for all study participants</td>
</tr>
<tr>
<td></td>
<td>d. Adequate description of the sampling frame and recruitment</td>
<td>d. Adequate description of participants lost to follow-up</td>
<td>d. The method and setting of measurement of PF is the same for all study participants</td>
<td></td>
</tr>
<tr>
<td></td>
<td>e. Adequate description of the period and place of recruitment</td>
<td>e. There are no important differences between participants who completed the study and those who did not</td>
<td>e. Adequate proportion of the study sample has complete data for the PF</td>
<td></td>
</tr>
<tr>
<td></td>
<td>f. Adequate description of inclusion and exclusion criteria</td>
<td>f. Appropriate methods of imputation are used for missing PF data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Study Confounding</td>
<td>6. Statistical Analysis and Reporting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Important potential confounding factors are appropriately accounted for</td>
<td>The statistical analysis is appropriate, and all primary outcomes are reported</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. All important confounders are measured</td>
<td>a. Sufficient presentation of data to assess the adequacy of the analytic strategy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Clear definitions of the important confounders measured are provided</td>
<td>b. Strategy for model building is appropriate and is based on a conceptual framework or model</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Measurement of all important confounders is adequately valid and reliable</td>
<td>c. The selected statistical model is adequate for the design of the study</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. The method and setting of confounding measurement are the same for all study participants</td>
<td>d. There is no selective reporting of results</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
e. Appropriate methods are used if imputation is used for missing confounder data

f. Important potential confounders are accounted for in the study design

g. Important potential confounders are accounted for in the analysis

| The observed effect of the PF on the outcome is very likely to be distorted by another factor related to PF and outcome | The reported results are very likely to be spurious or biased related to analysis or reporting |
| The observed effect of the PF on outcome may be distorted by another factor related to PF and outcome | The reported results may be spurious or biased related to analysis or reporting |
| The observed effect of the PF on outcome is unlikely to be distorted by another factor related to PF and outcome | The reported results are unlikely to be spurious or biased related to analysis or reporting |
Risk of bias assessment for genetic association studies

- Selection bias
- Information bias
- Confounding bias
- Multiple testing
- Selective reporting
- HWE

- Yes, low/no risk of bias; No, possible/high risk of bias; unclear

<table>
<thead>
<tr>
<th>Domain</th>
<th>Item</th>
<th>Low risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection bias</td>
<td>Representativeness of cases</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A. Consecutive/randomly selected from cases</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>population with clearly defined random frame</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>B. Consecutive/randomly selected from cases</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>population without clearly defined random frame or</td>
<td></td>
</tr>
<tr>
<td></td>
<td>with extensive inclusion criteria</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C. Spectrum of diseases</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Select on advance (atrophy or neovascular) or mild</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AMD</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A. Not describe method of selection</td>
<td></td>
</tr>
<tr>
<td>Representativeness of controls</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Controls were consecutive/randomly drawn from area (ward/community) as cases with the same criteria</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>B. Controls were consecutively/randomly drawn from different areas as cases</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>C. Not describe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Differential participation in case and control</td>
<td>Non-participant rate is small (&lt; 10%) and similar (to rates?) between case and control groups</td>
<td>Yes</td>
</tr>
<tr>
<td>Incomplete participant rates are different</td>
<td>Refusal or inability to provide data</td>
<td>NO</td>
</tr>
<tr>
<td></td>
<td>Refusal or inability to provide biological specimens</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Insufficient amount quality of data/ quality of DNA</td>
<td></td>
</tr>
<tr>
<td>Information bias</td>
<td>Ascertainment of AMD</td>
<td>Ascertainment of control</td>
</tr>
<tr>
<td>------------------</td>
<td>----------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td></td>
<td>Clearly described objective criteria of diagnosis of AMD</td>
<td>- Controls were non-case that proved by ocular examination</td>
</tr>
<tr>
<td></td>
<td>Not describe/unclear definition</td>
<td>- Just mentioned that controls were subjects who did not have AMD without ocular examination</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Confounding bias</td>
<td>Population stratification</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>- No difference in ethnic origin between cases and controls</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>- Use of controls who were not related to cases</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>- Use of some controls who came from the same family</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>- Other confounding controls</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>- Use of genomic controls</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>- Not reporting bias</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>- Controls for confounding variables (e.g., age, gender, smoking) in analysis</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>- Not controlled/not mentioned (or, no control/ no mention)</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Multiple testing &amp; Selective reporting (for replication studies)</th>
<th>How many polymorphisms have been studied</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Adjustment for multiple tests</td>
<td>Yes</td>
</tr>
<tr>
<td>- Report results of all polymorphisms mentioned in objectives, non-significant or not</td>
<td>Yes</td>
</tr>
<tr>
<td>- Report results of only significant polymorphisms</td>
<td>No</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HWE</th>
</tr>
</thead>
<tbody>
<tr>
<td>- HWE in control group</td>
</tr>
<tr>
<td>- HW disequilibrium in control group</td>
</tr>
<tr>
<td>- Not check HWE</td>
</tr>
</tbody>
</table>
Risk of bias assessment for prediction model

• PROBAST* jointly with CHARMS
• PROBAST comprises 5 domains
  • Participant selection
  • Outcome
  • Predictors
  • Sample size and missing data
  • Statistical analysis

Accessed 22/04/2018
Network

- Cochrane collaboration
- RCT
- Diagnostic studies
Welcome to The Cochrane Library

The Cochrane Library contains high-quality, independent evidence to inform healthcare decision-making. It includes reliable evidence from Cochrane and other systematic reviews, clinical trials, and more. Cochrane reviews bring you the combined results of the world's best medical research studies, and are recognised as the gold standard in evidence-based healthcare.

More About The Cochrane Library

What’s New in Issue 4, 2006?

109 new reviews, 57 updated reviews, 121 new protocols and 8 updated protocols, including:

- Screening for breast cancer with mammography
- Home-based chemically-induced whitening of teeth in adults
- Rimoban for overweight or obesity
- Stapled versus conventional surgery for hemorrhoids
- Regional versus general anaesthesia for caesarean section
- Soy formula for prevention of allergy and food intolerance in infants
- Occupational therapy for patients with problems in activities of daily living after stroke
- Interventions for increasing pedestrian and cyclist visibility for the prevention of death and injuries
- Progesterone for premenstrual syndrome

Access to The Cochrane Library

The Cochrane Library is available online through Wiley InterScience.

Help! New Users Start Here

As a new user we recommend you use the following resources to help you navigate through the evidence and get the most out of The Cochrane Library.

For Clinicians

As a clinician you are under constant pressure to have high-quality, up-to-date evidence at your fingertips.

For Researchers

The internet has given us instant access to a huge amount of research, but the large volume of available information is a problem in itself.

For Patients

Healthcare consumers and patients need high-quality evidence about the effectiveness of treatments.

For Policy Makers

As a policy maker or healthcare manager you are a generalist in search of high-quality information across a broad range of issues.
Welcome to HuGENet™

Human Genome Epidemiology Network, or HuGENet™ is a global collaboration of individuals and organizations committed to the assessment of the impact of human genome variation on population health and how genetic information can be used to improve health and prevent disease.

- Message from Dr. Muin Khoury

- Learn More About HuGENet™ Purpose, Goals and Activities
- Learn About HuGENet™ Coordinating Centers
- Learn How to Become a HuGENet™ Collaborator

HuGENet™ Menu

- What's New
- HuGE Funding
- HuGE Published Literature Db
- Genotype Prevalence Db
- HuGE Book
- HuGE Case Studies
- HuGE e-Journal Club
- HuGE Facts Sheets
- HuGE Reviews
- HuGE Workshops & Meetings
- HuGENet™ Collaborators
- Current HuGE articles
- HuGENet™ Publications
- Feedback
Renal Disease

- Cytokine gene polymorphisms and adverse outcomes in renal transplantation
  Ammarin Thakkinstian, Mark McEvoy, Steve Bowe, and John Attia
  (University of Newcastle, Australia)

Women’s Health

- Genetics of Preeclampsia
  Yves Giguere (Hôpital ST-François d’Assise, Quebec, Canada)

Other Health Associations

- Complement factor H polymorphism and age-related macular degeneration
  Ammarin Thakkinstian, Pearlne Han, Mark McEvoy, Wayne Smith, John Attia
  (University of Newcastle, Australia)

- Head and Neck
  Marko Lens (Imperial College London, UK)
  International Collaborative Study on Genetic Susceptibility to Environmental Carcinogens

- Y chromosome microdeletions in male infertility
  Borut Peterlin (UMC Ljubljana, Slovenia)

⚠️ Provides link to non-governmental sites and does not necessarily represent the views of the Centers for Disease Control and Prevention.
Register review proposal

Why do we need to register

• Establish that we are doing this review
• May reduce the risk of multiple reviews addressing the same question
• Increases potential communication with interested researchers
• Promote transparency of the methods
• Allows your peers to review how you will extract data for quantitative poolings
• Serve as a road map for our review
• What do we need in hands for registration
  • Research questions & specific objectives
  • Review methods,
    • How to identify studies
    • Selection of studies
    • Data extractions & risk of bias assessment
    • Interventions/Exposure
    • Outcomes of primary interest
    • Statistical analysis plan
    • Time schedule
Where to register

• National Institute of Health (NIH):
• Campbell Collaboration - produces systematic reviews of the
effects of social interventions
  • http://www.campbellcollaboration.org/
• Cochrane Collaboration - international organization,
produces and disseminates systematic reviews of health
care interventions
  • http://www.cochrane.org/
• PROSPERO - international prospective register of systematic reviews
  • http://www.crd.york.ac.uk/PROSPERO/
Grading the quality of evidence (GRADEpro)


The GRADE approach

• A method of grading the quality of evidence and strength of recommendations in health.

• Can be used to develop clinical practice guidelines and other health care recommendations.


• The GRADE working group is a collaboration of
  • methodologists, guideline developers, clinicians and other interested members

• Aim to developing and implementing a common, transparent and sensible approach to grading the quality of evidence and strength of recommendations in health care.
Purpose

• To develop Clinical Practice Guidelines (CPG)
• Offer recommendations for the management of typical patients.
• Management decisions involve balancing the desirable and undesirable effects of a given course of action.
• In order to help clinicians make evidence-based medical decisions, guideline developers often grade the strength of their recommendations.
Grading evidences

- Grade each pooling outcome
- Four rating scales

<table>
<thead>
<tr>
<th>Rating scale</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>We are highly confident that the true effect size lies close to our estimation. This means our pooled effect size is very precise estimation or 95% confidence interval (CI) is very narrow.</td>
</tr>
<tr>
<td>Moderate</td>
<td>We are moderately confident in the estimated effect size. The true effect may be close to the estimated effect size, but the 95% CI is quite wide and thus there is a possibility that it is substantially different.</td>
</tr>
<tr>
<td>Low</td>
<td>We are little confident in the estimated effect size because the estimation is very imprecise (i.e., very wide CI). The true effect may be substantially different from our estimated effect size.</td>
</tr>
<tr>
<td>Very low</td>
<td>We are very little confident in the estimated effect size, which has very wide CI: The true effect is likely to be substantially different from our estimate.</td>
</tr>
</tbody>
</table>
How to grade

Initial rating: Study design

• Randomized trials without important limitations provide high quality evidence
• Observational studies without special strengths or important limitations provide low quality evidence
• Non-randomised experimental trials (e.g., quasi-RCT) will automatically be downgraded for limitations in design, i.e., risk of bias in lack of concealment of allocation, blinding

• Case series and case reports
  • Investigate only patients exposed to the intervention.
  • Source of control group results is implicit or unclear, thus, they will usually warrant downgrading from low to very low quality evidence.
Grade diagram

Grading quality of evidence

What is the methodology of the best available evidence?

- randomised trial
  - assume high quality
  - high quality
    - factors lowering the quality present?
      - NO
        - moderate, low, or very low quality
      - YES
        - downgrade depending on the total number and weight of limitations
      - upgrade depending on the total number and weight of additional merits
  - moderate, low, or very low quality
    - upgrade depending on the total number and weight of additional merits
  - high, moderate or low quality
- observational study
  - assume low quality
  - factors lowering the quality present?
    - NO
      - low quality
    - YES
      - very low quality
  - very low quality
    - upgrade depending on the total number and weight of additional merits
  - moderate or high quality

Wisdom of the Land
Five factors are used to possibly rate down the quality of evidence

<table>
<thead>
<tr>
<th>Down-rate factor</th>
<th>Consequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limitations in study design or execution (risk of bias)</td>
<td>↓ 1 or 2 levels</td>
</tr>
<tr>
<td>Inconsistency of results</td>
<td>↓ 1 or 2 levels</td>
</tr>
<tr>
<td>Indirectness of evidence</td>
<td>↓ 1 or 2 levels</td>
</tr>
<tr>
<td>Imprecision</td>
<td>↓ 1 or 2 levels</td>
</tr>
<tr>
<td>Publication bias</td>
<td>↓ 1 or 2 levels</td>
</tr>
<tr>
<td>Up-rate factor</td>
<td></td>
</tr>
<tr>
<td>Large magnitude of effect</td>
<td>↑ 1 or 2 levels</td>
</tr>
<tr>
<td>All plausible confounding would reduce the demonstrated effect or increase the</td>
<td>↑ 1 level</td>
</tr>
<tr>
<td>effect if no effect was observed</td>
<td></td>
</tr>
<tr>
<td>Dose-response gradient</td>
<td>↑ 1 level</td>
</tr>
</tbody>
</table>
1. Risk of bias

- Confidence in the estimated effect decreases if studies suffer from major limitations that are likely to result in a biased assessment of the intervention effect
- RBA tool for RCTs
<table>
<thead>
<tr>
<th>sequence generation</th>
<th>The patients are assigned to either the treatment or control group through a process not based entirely on chance and can be predicted (major problem in “pseudo” or “quasi” randomized trials with allocation by day of week, birth date, chart number etc.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of allocation concealment</td>
<td>Those enrolling patients are aware of the group to which the next enrolled patient will be allocated</td>
</tr>
<tr>
<td>Lack of blinding</td>
<td>Patient, caregivers, those recording outcomes, those adjudicating outcomes, or data analysts are aware of the arm to which patients are allocated</td>
</tr>
<tr>
<td>Incomplete accounting of patients and outcome events</td>
<td>Loss to follow-up and failure to adhere to the intention to treat principle when indicated. The significance of particular rates of loss to follow-up, however, varies widely and is dependent on the relation between loss to follow-up and number of events</td>
</tr>
<tr>
<td>Selective outcome reporting</td>
<td>Reporting of some outcomes and not others on the basis of the results. For example, review authors would need the protocol for the study to be certain this bias is not present.</td>
</tr>
<tr>
<td>Other limitations</td>
<td>For example: • stopping early for benefit observed in randomized trials, in particular in the absence of adequate stopping rules (tends to overestimate treatment effect in trials with fewer than 500 events) • use of unvalidated patient-reported outcomes • carry-over effects in cross-over trials • recruitment bias in cluster-randomized trials</td>
</tr>
<tr>
<td>Limitation</td>
<td>explanation</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Failure to develop and apply appropriate eligibility criteria (inclusion</td>
<td>• under- or over-matching in case-control studies</td>
</tr>
<tr>
<td>of control population)</td>
<td>• selection of exposed and unexposed in cohort studies from different</td>
</tr>
<tr>
<td></td>
<td>populations</td>
</tr>
<tr>
<td></td>
<td>• differences in measurement of exposure (e.g. recall bias in case-control</td>
</tr>
<tr>
<td></td>
<td>studies)</td>
</tr>
<tr>
<td></td>
<td>• differential surveillance for outcome in exposed and unexposed in cohort</td>
</tr>
<tr>
<td></td>
<td>studies</td>
</tr>
<tr>
<td>Flawed measurement of both exposure and outcome</td>
<td>• failure of accurate measurement of all known prognostic factors</td>
</tr>
<tr>
<td></td>
<td>• failure to match for prognostic factors and/or adjustment in statistical</td>
</tr>
<tr>
<td></td>
<td>analysis</td>
</tr>
<tr>
<td>Failure to adequately control confounding</td>
<td>• especially within prospective cohort studies, both groups should be</td>
</tr>
<tr>
<td></td>
<td>followed for the same amount of time</td>
</tr>
<tr>
<td>Incomplete or inadequately short follow-up</td>
<td></td>
</tr>
</tbody>
</table>
From RBA to judgements about study limitations for outcomes

<table>
<thead>
<tr>
<th>Risk of bias</th>
<th>Across studies</th>
<th>Interpretation</th>
<th>Considerations</th>
<th>GRADE assessment of study limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Most information is from studies at low risk of bias.</td>
<td>Plausible bias unlikely to seriously alter the results.</td>
<td>No apparent limitations.</td>
<td>No serious limitations, do not downgrade.</td>
</tr>
<tr>
<td>Unclear</td>
<td>Most information is from studies at low or unclear risk of bias.</td>
<td>Plausible bias that raises some doubt about the results.</td>
<td>Potential limitations are unlikely to lower confidence in the estimate of effect.</td>
<td>No serious limitations, do not downgrade.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Potential limitations are likely to lower confidence in the estimate of effect.</td>
</tr>
<tr>
<td>High</td>
<td>The proportion of information from studies at high risk of bias is sufficient to affect the interpretation of results.</td>
<td>Plausible bias that seriously weakens confidence in the results.</td>
<td>Crucial limitation for one criterion, or some limitations for multiple criteria, sufficient to lower confidence in the estimate of effect.</td>
<td>Serious limitations, downgrade one level.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Crucial limitation for one or more criteria sufficient to substantially lower confidence in the estimate of effect.</td>
</tr>
</tbody>
</table>
2. Inconsistency of Results

• If treatment effects are varied across studies,
  • Investigators should explore explanations for heterogeneity, and
  • If they cannot identify a plausible explanation, the quality of evidence should be downgraded.
  • Whether it is downgraded by one or two levels will depend on the magnitude of the inconsistency in the results.
    • Moderate $I^2$ but can explain source/s of heterogeneity, not downgrade
    • Moderate $I^2$ but cannot explain source/s of heterogeneity, down 1 grade
    • High $I^2$ with explaining source/s of heterogeneity, not down or down 1 grade
    • High $I^2$ without explaining source/s of heterogeneity, down 2 grade

• Sources of heterogeneity
  • Populations (e.g. drugs may have larger relative effects in sicker populations)
  • Interventions (e.g. larger effects with higher drug doses)
  • Outcomes (e.g. diminishing treatment effect with time)
  • Study methods (e.g. differences in the risk of bias).
3. Indirectness of Evidence

• Different in targeted patients
  • Studies had differences in patient spectrum
  • Disease severity
  • Mixed population of children and adults,
  • Specific disease/s, e.g., diabetes or hypertension but results will be used in general adults
  • This evidence is said to be indirectness.

• Difference in intervention
  • Different in interested intervention, drug class, or dosage per day between studies, can be used for down grade the evidence.
• Difference in the outcome measures
  • Different methods used for outcome measure
  • Different time at measure
  • Surrogate versus patient important outcome.
    • Using surrogate outcome can be down grade by one or two level. Knowing disease mechanism will help in making decision to down grade, if the surrogate outcome is far away from the patient important outcome, it should be down grade by two levels, e.g.,
      • Bone mineral density for fracture, calcium calcification score for myocardial infraction should be down grade by one level;
      • Sleep questionnaire to measure sleep quality in pregnancy me be down grade by two levels.
• Indirect comparison
  • If treatment effect C vs B is indirectly estimated from B vs A and C vs A, indirect effect of C vs B should be down grade by one level, and two levels if inconsistency assumption is violated.
4. Imprecision of the outcome

• The optimal information size (OIS) is used to determine whether 95% CI for each outcome is adequate precision.

• The OIS is a **total number of patients included in a systematic review**

• If the OIS is lower than a conventional sample size calculation n), rate down for imprecision, unless the sample size is very large (at least 2000, and perhaps 4000 patients).
• If the OIS criterion is met (OIS ≥ conventional n) and the **95% CI excludes no effect** (i.e. CI around RR excludes 1.0), **do not rate down** for imprecision.

• If OIS criterion is met, and the **95% CI overlaps no effect** (i.e. CI includes RR of 1.0) **rate down for imprecision** if the CI **fails to exclude important benefit or important harm**.
  
  • The lower CI is very benefit or the upper CI is very high risk, it may rate down for one level.
  
  • If the lower CI is very benefit but the upper CI is high risk, it may consider to rate down for two levels.
Upgrading quality of evidence

Large Magnitude of the Effect

- The larger the magnitude of effect, the stronger becomes the evidence.
- Only studies with no threats to validity (not downgraded for any reason) can be upgraded.

Evidences for

- Observational studies with strong methodology (e.g. with adjusted analyses)
- RCTs
- RCTs in which randomisation is compromised (e.g. inadequate sequence generation) but the analyses have been adjusted, could also be upgraded
Upgrading

- ORs can be used with this condition if disease/event is < 0.2
- If disease/event is more common (> 0.40), then RR is larger than OR, thus need larger threshold

<table>
<thead>
<tr>
<th>Magnitude of effect</th>
<th>Effect measure</th>
<th>Quality of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>large</td>
<td>RR &gt;2 or &lt;0.5 (based on consistent evidence from at least 2 studies, with no plausible confounders)</td>
<td>upgrade 1 level</td>
</tr>
<tr>
<td>very large</td>
<td>RR &gt;5 or &lt;0.2 (based on direct evidence with no major threats to validity)</td>
<td>upgrade 2 levels</td>
</tr>
</tbody>
</table>
The Effect of all Plausible Confounding

• All plausible confounding from observational studies or RCTs may reduce the demonstrated effect or increase the effect if no effect was observed.

• Studies with no important threats to validity should be upgraded.
  • Methodologically strong observational studies (e.g. with adjusted analyses) or, in theory,
  • RCTs with strong methodology
  • RCTs in which randomisation is compromised (e.g. inadequate sequence generation) but the analyses have been adjusted
Dose-Response Gradient

• A dose-response relationship lends supporting a possible cause-and-effect relationship
• The presence of a dose-response gradient may increase our confidence in the findings of observational studies and thereby increase the quality of evidence.
• Only studies with no threats to validity (not downgraded for any reason) can be upgraded.