CRITICAL APPRAISAL OF MEDICAL LITERATURE

Samuel Iff
ISPM Bern
siff@ispm.unibe.ch
Contents

• Study designs
• Asking good questions
• Pitfalls in clinical studies
• How to assess validity (RCT)
• Conclusion
Step-by-step

- What is my question I want answered (PICO)
- What study type is best to get that answer (Study type)
- Find the paper (literature search)
- Was the study done properly (Internal validity)
- Does it apply to my patients (External validity)
STUDY DESIGN

Repetition
Study types

Medical research

Primary research

Basic research
- Theoretical
- Method development (physics, chemistry, biology, bioinformatics, biometrics, psychology)
- Analytical measurement procedure
- Imaging procedure
- Biometric procedure
- Test development

Clinical research
- Experimental
- Clinical study
- Phase I study
- Phase II study
- Phase III study
- Phase IV study

Epidemiological research
- Observational

Secondary research

Meta-analysis
- Review
- Systematic
- Simple (narrative)

Evidence
Classification of evidence

- Systematic reviews
- Randomized Controlled Trials
- Cohort studies
- Case-Control Studies
- Case-Series, Case Reports
- Descriptive study, Editorial, Expert Opinion
## Classification of evidence

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level I:</td>
<td>Evidence obtained from at least one properly designed randomized controlled trial or meta-analysis.</td>
</tr>
<tr>
<td>Level II-1:</td>
<td>Evidence obtained from well-designed controlled trials without randomization.</td>
</tr>
<tr>
<td>Level II-2:</td>
<td>Evidence obtained from well-designed cohort or case-control analytic studies, preferably from more than one center or research group.</td>
</tr>
<tr>
<td>Level II-3:</td>
<td>Evidence obtained from multiple time series with or without the intervention. Dramatic results in uncontrolled trials might also be regarded as this type of evidence.</td>
</tr>
<tr>
<td>Level III:</td>
<td>Opinions of respected authorities, based on clinical experience, descriptive studies, or reports of expert committees.</td>
</tr>
</tbody>
</table>
Cross-sectional study

Population of interest → Representative sample → Measurement of characteristics
Cross-sectional study

Advantages

• can obtain findings quickly
• can often be undertaken with minimal funding
• multiple outcomes

Disadvantages

• Prevalence only
• No cause and effect association possible
• Confounding
Case-control study

Study population

Outcome yes

Exposure yes

Exposure no

Outcome no

Exposure yes

Exposure no
# Case-control study

## Advantages
- can obtain findings quickly
- can often be undertaken with minimal funding
- efficient for rare diseases
- can study multiple exposures
- generally requires few study subjects

## Disadvantages
- subject to bias (sampling bias, observation and recall bias)
- difficult if record keeping is either inadequate or unreliable
- selection of controls can be difficult
**Cohort study**

- **Study population**
  - **Exposure**
    - **No exposure**
      - Outcome yes
      - Outcome no
  - Outcome yes
  - Outcome no
Cohort study

Advantages
• can replace RCTs in unethical settings
• cause-effect relationship is measurable
• examine various outcomes
• effect of each variable on outcome is measurable
• retrospective is cheap

Disadvantages
• follow-up might be difficult
• recall bias in retrospective setting
• confounding variables
Randomized controlled trial

- Study population
- Randomization
  - Intervention
    - Outcome yes
    - Outcome no
  - Control
    - Outcome no
    - Outcome yes
Randomized controlled trial

Advantages

- “Gold standard”

Disadvantages

- Bias
- Expensive
Study type

- The study type defines the level of evidence
- Each study is subjected to its own advantages and weaknesses
- Each clinical question has an optimal study design.
ASKING GOOD QUESTIONS
## Clinical questions

| The Swiss government needs to provide figures to WHO on the incidence of prostate cancer in Switzerland | Annas mum is worried about Anna using her mobile phone so much – she heard that they are not safe |
| Mrs. Smiths GP is wondering whether acupuncture might help Mrs. Smiths shoulder pain | Lara has a positive HIV-test but was never exposed to a risk |
Medical question

- What are the problems? (Observation)
- How common is it? (Frequency)
- What caused it? (Aetiology)
- Who has the problem? (Diagnosis)
- What will happen? (Prognosis)
- How will the intervention change the outcome? (Intervention)
## Question and study design

<table>
<thead>
<tr>
<th>Question</th>
<th>Possible study design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention/ Outcome</td>
<td>Randomized controlled Trial</td>
</tr>
<tr>
<td></td>
<td>Cohort study</td>
</tr>
<tr>
<td></td>
<td>Case-control-study</td>
</tr>
<tr>
<td></td>
<td>Ecological study</td>
</tr>
<tr>
<td></td>
<td>Pre-Post-study</td>
</tr>
<tr>
<td>Aetiology</td>
<td>Randomized controlled Trial</td>
</tr>
<tr>
<td></td>
<td>Cohort study</td>
</tr>
<tr>
<td></td>
<td>Case-control-study</td>
</tr>
<tr>
<td></td>
<td>Ecological study</td>
</tr>
<tr>
<td></td>
<td>Pre-Post-study</td>
</tr>
<tr>
<td>Diagnosis (Test vs. goldstandard)</td>
<td>Cross-sectional study</td>
</tr>
<tr>
<td>Diagnosis (Comparison of tests)</td>
<td>Randomized controlled Trial</td>
</tr>
<tr>
<td></td>
<td>Cohort study</td>
</tr>
<tr>
<td></td>
<td>Case-control-study</td>
</tr>
<tr>
<td>Incidence</td>
<td>Descriptive cross-sectional study</td>
</tr>
<tr>
<td></td>
<td>Descriptive cohort study</td>
</tr>
<tr>
<td>Prognosis/ Frequency</td>
<td>Cohort study</td>
</tr>
<tr>
<td><strong>PICO</strong></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td></td>
</tr>
<tr>
<td><strong>Population</strong></td>
<td>Who are the relevant patients and what is the problem</td>
</tr>
<tr>
<td><strong>Intervention</strong></td>
<td>What is the treatment being considered</td>
</tr>
<tr>
<td><strong>Comparator</strong></td>
<td>What is it compared to</td>
</tr>
<tr>
<td><strong>Outcome</strong></td>
<td>What are the person-relevant consequences of the exposure</td>
</tr>
</tbody>
</table>
Example

- Mrs. Smiths GP is wondering whether accupuncture might help Mrs. Smiths shoulder pain

- **PICO**
  - In patients with chronic shoulder pain…
  - … is accupuncture…
  - … compared to placebo…
  - … reducing pain?
Exercise 1

• PICO:

• You are referred a patient with recurrent debilitating migraine headaches, who has tried several prophylactic treatments. He comes to clinic with an alternative health magazine containing an article which claims a breakthrough in migraine control using acupuncture, and print-outs from several websites and asks your opinion about whether you acupuncture will lessen his attacks. You are unsure whether acupuncture can be recommended.
**Answer 1**

<table>
<thead>
<tr>
<th>P</th>
<th>In people with migraine..</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>.. does accupuncture..</td>
</tr>
<tr>
<td>C</td>
<td>.. compared to regular medication..</td>
</tr>
<tr>
<td>O</td>
<td>.. lessen the frequency or severity of attacks?</td>
</tr>
<tr>
<td>Study-Type</td>
<td>Intervention</td>
</tr>
</tbody>
</table>
PICO-S

• Patients
• Intervention
• Comparator
• Outcome
• Study type

• Asking the right question will lead you to the right answer. Only good answers are useful in epidemiology.
PITFALLS IN CLINICAL STUDIES
Bias

- Systematic difference between observed result and the truth
Selection Biases

Selection biases occur when the groups to be compared are different. These differences may influence the outcome.

- **Volunteer or referral bias**: Volunteer or referral bias occurs because people who volunteer to participate in a study (or who are referred to it) are often different than non-volunteers/non-referrals. This bias usually, but not always, favors the treatment group, as volunteers tend to be more motivated and concerned about their health.

- **Non-respondent bias**: Non-respondent bias occurs when those who do not respond to a survey differ in important ways from those who respond or participate. This bias can work in either direction.
Measurement Biases

Measurement biases involve systematic error that can occur in collecting relevant data.

- **Instrument bias.** Instrument bias occurs when calibration errors lead to inaccurate measurements being recorded, e.g., an unbalanced weight scale.
- **Insensitive measure bias.** Insensitive measure bias occurs when the measurement tool(s) used are not sensitive enough to detect what might be important differences in the variable of interest.
- **Expectation bias.** Expectation bias occurs in the absence of masking or blinding, when observers may err in measuring data toward the expected outcome. This bias usually favours the treatment group.
- **Recall or memory bias.** Recall or memory bias can be a problem if outcomes being measured require that subjects recall past events. Often a person recalls positive events more than negative ones. Alternatively, certain subjects may be questioned more vigorously than others, thereby improving their recollections.
- **Attention bias.** Attention bias occurs because people who are part of a study are usually aware of their involvement, and as a result of the attention received may give more favourable responses or perform better than people who are unaware of the study’s intent.
- **Verification or work-up bias.** Verification or work-up bias is associated mainly with test validation studies. In these cases, if the sample used to assess a measurement tool (e.g., diagnostic test) is restricted only to who have the condition of factor being measured, the sensitivity of the measure can be overestimated.
Intervention (Exposure) Biases

Intervention or exposure biases generally are associated with research that compares groups.

- **Contamination bias.** Contamination bias occurs when members of the 'control' group inadvertently receive the treatment or are exposed to the intervention, thus potentially minimizing the difference in outcomes between the two groups.

- **Co-intervention bias.** Co-intervention bias occurs when some subjects are receiving other (unaccounted for) interventions at the same time as the study treatment.

- **Timing bias(es).** Different issues related to the timing of intervention can bias. If an intervention is provided over a long period of time, maturation alone could be the cause for improvement. If treatment is very short in duration, there may not have been sufficient time for a noticeable effect in the outcomes of interest.

- **Compliance bias.** Compliance bias occurs when differences in subject adherence to the planned treatment regimen or intervention affect the study outcomes.

- **Withdrawal bias.** Withdrawal bias occurs when subjects who leave the study (drop-outs) differ significantly from those that remain.

- **Proficiency bias.** Proficiency bias occurs when the interventions or treatments are not applied equally to subjects. This may be due to skill or training differences among personnel and/or differences in resources or procedures used at different
Bias versus chance

![Graph showing bias versus chance in diastolic blood pressure measurements.](image)
HOW TO ASSESS VALIDITY
Can I use the results for my patients
Bias in RCTs

- Patients
- Randomization
  - Intervention
  - Control
    - Outcome assessment
    - Outcome assessment
    - Analysis
    - Analysis
- Selection bias
  - Performance bias
  - Detection bias
  - Attrition bias
Bias and its remedy

- Selection bias
- Performance bias
- Detection bias
- Attrition bias
- Concealment of allocation
- Blinding of patients/doctors
- Blinding of assessors
- Intention-to-treat analysis
Internal validity

- **Selection bias** - Systematic error in creating intervention groups, such that they differ with respect to prognosis.
- **Performance bias** - Systematic differences in the care provided to the participants in the comparison groups other than the intervention under investigation.
- **Detection bias** - Systematic difference between comparison groups in how outcomes are ascertained, diagnosed or verified.
- **Attrition bias** - Systematic differences between comparison groups in withdrawals or exclusions of participants from the results of a study.

- Careful design, implementation and analysis lead to good results
Anticoagulation for Myocardial Infarction

Myocardial Infarction

Allocation according to entry dates

Odd Dates
Anticoagulants

Even Dates
Conventional

Wright et al, Am Heart J 1948
Anticoagulation for Myocardial Infarction

Odd Dates
- Anticoagulants
- 589 Patients

Even Dates
- Conventional
- 442 Patients

P=0.001

Generation of allocation sequence

• Adequate if resulting sequences are unpredictable:
  • coin tossing
  • computer generated random-numbers
  • drawing lots or envelopes
  • shuffling cards
  • table of random-numbers
  • throwing dice

• Inadequate if resulting sequences are predictable:
  • according to case record number
  • according to date of birth
  • according to date of admission
  • alternation
Concealment of allocation

Adequate if patients and enrolling physicians cannot foresee assignment:

- central randomisation
- a priori and independently numbered or coded drug packs
- sequentially numbered, sealed, opaque envelopes

Inadequate if patients or enrolling physicians can foresee assignment:

- procedures based on inadequate generation of sequences
- open allocation schedule
- unsealed or non-opaque envelopes
Plasmapheresis for Multiple Sclerosis

Plasmapheresis

Sham

Assessment by
1 blinded and 1 unblinded neurologist

Change in Expanded Disability Status Scale at 6, 12, 18 and 24+ months

Noseworthy et al, Neurology 1994
Plasmapheresis for Multiple Sclerosis

Noseworthy et al, Neurology 1994

Sham - treating physicians
Plasmapheresis - treating physicians
Sham - blinded assessors
Plasmapheresis - blinded assessors

Median change EDSS

Time from randomization

0 1 2 years

Noseworthy et al, Neurology 1994
### Plasmapheresis for Multiple Sclerosis

<table>
<thead>
<tr>
<th></th>
<th>Blinded</th>
<th>Unblinded</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 months</td>
<td>0.246</td>
<td>0.047</td>
</tr>
<tr>
<td>12 months</td>
<td>0.086</td>
<td>0.004</td>
</tr>
<tr>
<td>18 months</td>
<td>0.106</td>
<td>0.072</td>
</tr>
<tr>
<td>24 months</td>
<td>0.201</td>
<td>0.031</td>
</tr>
</tbody>
</table>

*p values for between tx comparison of proportion of subjects improved, stable, or worse*
Surgical Therapy for Bilateral Carotid Stenosis

TIA and bilateral carotid stenosis

Surgical

Medical

Analysis

Available patients

Intention to treat

Surgical Therapy for Bilateral Carotid Stenosis

Patients available for follow-up
(surgical = 79, medical = 72)

OR 0.43 (0.20, 0.90)

Outcome: Recurrent TIA, stroke or death
Intention-to-treat analysis

• Analysis on the basis of the comparison arm to which participant was randomised
• Not on the basis of
  • Post-randomisation assessment of eligibility
  • Whether treatment was received
  • Whether treatment was completed
CONCLUSION
PICO-SU

- Patients
- Intervention
- Comparator
- Outcome
- Study type

- Form a clinically relevant question
Internal validity (RCT)

- Selection bias
- Performance bias
- Detection bias
- Attrition bias
External validity

Can the results be applied to my patients
• Study patients are comparable to my patients
• Intervention is the same as in the study
• External factors are the same (e.g. race, social status, resources, medical system)
• Same outcome is important in the study as in my patients (e.g. same follow-up time, relevant result)
• Clinically relevant end-points
• Clinically relevant effect size
Stay critical

- Ask for the best evidence
- Question the results
- Use your brain
- Never trust a statistic you did not forge yourself 😊
THANKS 😊