

Mastering Critical Appraisal in Medicine

Enhance Your Skills in Evaluating Clinical Evidence with EviScore

EviScore Foundation Course

Module 3: APPRAISE

Critically Evaluating Clinical Evidence

Module Learning Journey

Module 3: APPRAISE

↓

Chapter 1: Introduction to Critical Appraisal

↓

Chapter 2: Appraising Therapy Studies

↓

Chapter 3: Appraising Diagnostic, Prognostic & Harm Studies

↓

Chapter 4: Appraising Systematic Reviews & Guidelines

↓

Final Knowledge Check

Learning Outcomes

By the end of this module, you will be able to:

- **Explain** the purpose of critical appraisal
 - **Describe** the three fundamental questions of critical appraisal
 - **Recognize** common sources of bias
 - **Distinguish** random error from systematic error
 - **Understand** why validity must be assessed before applying research findings
 - **Apply** appropriate appraisal frameworks to different study designs
-

Why This Module Matters

Finding a research article does not mean its conclusions should be accepted.

Evidence-Based Healthcare requires clinicians to determine whether evidence is **valid**, **important**, and **applicable** before using it in patient care.

EVIDENCE PEARL

Not all published evidence is trustworthy. Critical appraisal helps clinicians separate reliable evidence from misleading evidence.

CHAPTER 1

Introduction to Critical Appraisal

What is Critical Appraisal?

Critical appraisal is the **systematic evaluation** of research evidence to determine:

- Whether the study was conducted appropriately
- Whether the results are trustworthy
- Whether the findings are relevant to patient care

Critical Appraisal Is NOT:

- ✗ Finding faults in every paper
- ✗ Rejecting studies because they are imperfect
- ✗ Accepting conclusions without evaluation

Critical Appraisal IS:

- ✓ Asking structured questions
- ✓ Judging study quality
- ✓ Assessing the strength of evidence
- ✓ Making informed clinical decisions

EVIDENCE PEARL

A published article is evidence, not necessarily good evidence.

The Three Fundamental Questions

Every critical appraisal follows the same framework, regardless of study design:

1. Are the Results Valid?

- Was the study designed appropriately?

- Was bias minimized?
- Can the results be trusted?

2. What Are the Results?

- What was the magnitude of the effect?
- How precise are the estimates?
- Are the findings clinically meaningful?

3. Can I Apply the Results?

- Are my patients similar to those studied?
- Were important outcomes measured?
- Do the benefits outweigh harms and costs?

Universal Framework

Validity → Results → Applicability

EVIDENCE PEARL

Always evaluate validity first. If a study is seriously flawed, interpreting its results may be misleading.

Understanding Bias

What Is Bias?

Bias is a systematic error in the design, conduct, analysis, or reporting of a study that leads to an incorrect estimate of the true effect.

Unlike random error, bias **consistently** moves the results away from the truth.

Common Sources of Bias:

- **Selection bias** - Problems with how participants were chosen
- **Performance bias** - Differences in care between groups
- **Detection bias** - Differences in outcome assessment
- **Attrition bias** - Loss of participants during follow-up
- **Reporting bias** - Selective reporting of results

Why Does Bias Matter?

Bias can:

- Exaggerate treatment benefits
- Underestimate harms
- Create misleading conclusions

EVIDENCE PEARL

The purpose of critical appraisal is not to eliminate all bias—it is to recognize its presence and judge its impact.

Random Error vs Systematic Error

Understanding this distinction is essential when interpreting research.

Random Error (Chance)

Occurs because of natural variation

Characteristics:

- Unpredictable
- Reduced by larger sample sizes
- Reflected in confidence intervals

Example: A small randomized trial reports inconsistent treatment effects because relatively few participants were enrolled.

Systematic Error (Bias)

Occurs because of flaws in study design or conduct

Characteristics:

- Predictable distortion
- Not corrected by increasing sample size
- Threatens study validity

Example: Participants in one treatment group are substantially healthier at baseline than those in the comparison group.

Comparison Table

Aspect	Random Error	Systematic Error
Due to	Chance	Bias
Reduced by	Larger samples	Better study design
Affects	Precision	Validity

EVIDENCE PEARL

Large studies can still produce misleading results if they are systematically biased.

CHAPTER TAKEAWAY

- ✓ **Understand** that critical appraisal is essential for evidence-based practice
 - ✓ **Recognize** the three fundamental questions: validity, results, applicability
 - ✓ **Distinguish** between bias (systematic error) and chance (random error)
 - ✓ **Remember** that precision cannot compensate for bias
-

CHAPTER 2

Appraising Therapy Studies

Why Therapy Studies Matter

Many clinical decisions involve choosing between treatments:

- Drug A versus Drug B
- Surgery versus conservative treatment
- Exercise therapy versus usual care
- Intensive rehabilitation versus standard rehabilitation

Randomized Controlled Trials (RCTs) are the preferred design for therapy questions because randomization aims to create groups with similar prognoses before treatment begins.

Purpose of Therapy Appraisal

To determine whether the reported treatment effect is:

- ✓ **Valid**
- ✓ **Clinically important**
- ✓ **Applicable to your patient**

EVIDENCE PEARL

Not every randomized trial provides trustworthy evidence.

Step 1: Assess the Risk of Bias

Did the Study Start Fairly?

Key Questions:

Were patients randomized?

Randomization reduces selection bias by giving participants an equal chance of receiving each intervention.

Was allocation concealed?

Investigators enrolling participants should not know the upcoming treatment assignment.

Were the groups similar at baseline?

Important characteristics such as age, disease severity, and prognostic factors should be comparable.

Why Are These Questions Important?

Differences between groups before treatment begins may influence outcomes independently of the intervention.

EVIDENCE PEARL

Randomization creates comparable groups; allocation concealment protects the randomization process.

Step 2: Was Prognostic Balance Maintained?

Even if groups begin equally, bias may occur during the trial.

Questions to Ask:

Were participants, clinicians, and outcome assessors blinded?

Blinding reduces the risk that expectations influence treatment or outcome assessment.

Was follow-up complete?

Large losses to follow-up may distort the results.

Were patients analysed in the groups to which they were randomized?

This is known as **intention-to-treat analysis**.

Was the trial stopped early?

Trials stopped prematurely may overestimate treatment benefits.

EVIDENCE PEARL

A fair trial should remain fair from the beginning to the end.

Step 3: What Are the Results?

Once study validity has been assessed, evaluate:

Magnitude of Effect

How large is the difference between groups?

Common measures include:

- Relative Risk (RR)
- Absolute Risk Reduction (ARR)
- Risk Difference
- Number Needed to Treat (NNT)

Precision

How certain is the estimate?

Precision is commonly expressed using **95% confidence intervals**.

Interpretation

Large effect + Narrow confidence interval = Greater confidence in the estimate

EVIDENCE PEARL

Always interpret the effect size together with its confidence interval.

Step 4: Can I Apply These Results?

A valid study does not automatically apply to every patient.

Questions to Consider:

Were the study participants similar to my patient?

Consider:

- Age
- Disease severity
- Comorbidities
- Healthcare setting

Were all patient-important outcomes measured?

Examples:

- Survival
- Functional independence
- Quality of life
- Adverse events

Do the likely benefits outweigh the harms and costs?

Clinical decisions require balancing effectiveness with safety, patient preferences, and resource use.

EVIDENCE PEARL

Evidence becomes clinical care only when it fits the individual patient.

CHAPTER TAKEAWAY

- ✓ **Apply** the validity-results-applicability framework to therapy studies
 - ✓ **Recognize** key sources of bias in randomized trials
 - ✓ **Understand** that both effect size and precision matter
 - ✓ **Remember** that valid evidence must still be applicable to your patient
-

CHAPTER 3

Appraising Diagnostic, Prognostic & Harm Studies

Although these study designs answer different clinical questions, they share the same appraisal framework.

Appraising Diagnostic Studies

Why Diagnostic Studies Matter

Before treating a patient, clinicians must determine whether the patient truly has the disease.

Diagnostic studies evaluate how accurately a test distinguishes patients **with** disease from those **without** disease.

Key Validity Questions:

- ✓ Was there an independent blind comparison with the reference standard?
- ✓ Did the study include an appropriate spectrum of patients?
- ✓ Was every patient evaluated using both the index test and the reference standard?
- ✓ Were test methods described clearly enough for replication?

Important Measures:

- **Sensitivity** - How well the test identifies disease when present
- **Specificity** - How well the test excludes disease when absent
- **Positive Predictive Value** - Probability of disease when test is positive
- **Negative Predictive Value** - Probability of no disease when test is negative
- **Likelihood Ratios** - How much the test changes disease probability

EVIDENCE PEARL

A diagnostic test is only as trustworthy as the study evaluating it.

Appraising Prognosis Studies

Why Prognosis Matters

Patients often ask:

- What is likely to happen?
- Will I recover?
- What is my risk of complications?
- How long will recovery take?

Prognosis studies help answer these questions by **following patients over time**.

Key Validity Questions:

- ✓ Was there a representative and well-defined patient sample?
- ✓ Were patients enrolled at a similar stage of disease?
- ✓ Was follow-up sufficiently long and complete?
- ✓ Were outcomes measured objectively?
- ✓ Were important prognostic factors considered?

Common Outcomes:

- Survival
- Functional recovery
- Recurrence
- Disability
- Mortality

EVIDENCE PEARL

Good prognosis studies begin with comparable patients and complete follow-up.

Appraising Harm Studies

Why Harm Studies Matter

Clinicians must determine whether an exposure increases the risk of adverse outcomes.

Examples:

- Smoking and lung cancer
- Steroid use and osteoporosis
- Obesity and osteoarthritis
- Air pollution and respiratory disease

Key Validity Questions:

- ✓ Were comparison groups clearly identified?
- ✓ Were groups similar apart from the exposure?
- ✓ Were exposure and outcomes measured consistently?
- ✓ Was follow-up sufficiently complete?

- ✓ Was the temporal relationship appropriate?
- ✓ Was there evidence of a dose-response relationship?

Common Measures:

- **Relative Risk** - How much the exposure increases risk
- **Odds Ratio** - Odds of outcome in exposed vs unexposed
- **Number Needed to Harm** - How many need exposure for one harm

EVIDENCE PEARL

Association does not automatically imply causation.

Universal Framework Comparison

Study Type	Main Question	Preferred Design
Diagnosis	Is the test accurate?	Diagnostic Accuracy Study
Prognosis	What will happen?	Cohort Study
Harm	Does exposure increase risk?	Cohort/Case-Control Study

Universal Appraisal Framework:

Validity → Results → Applicability

This framework should guide every critical appraisal regardless of study design.

EVIDENCE PEARL

The study design changes, but the clinician's appraisal approach remains consistent.

CHAPTER TAKEAWAY

- ✓ **Understand** that different study designs answer different clinical questions
 - ✓ **Apply** the same three-question framework across all study types
 - ✓ **Recognize** design-specific sources of bias
 - ✓ **Remember** that validity, results, and applicability always matter
-

CHAPTER 4

Appraising Systematic Reviews & Guidelines

Appraising Systematic Reviews

Why Systematic Reviews Matter

A **systematic review** summarizes the results of multiple studies that address the same focused clinical question using explicit and reproducible methods.

Rather than relying on a single study, systematic reviews provide a **broader view** of the available evidence.

Key Validity Questions:

- ✓ Did the review address a focused clinical question?
- ✓ Were inclusion and exclusion criteria appropriate?
- ✓ Was the literature search comprehensive?
- ✓ Was the quality of included studies assessed?
- ✓ Were study assessments reproducible?
- ✓ Were the study results sufficiently similar?

Questions About Results:

- ✓ What are the overall results?
- ✓ Are the findings consistent across studies?
- ✓ How precise are the pooled estimates?

Applicability Questions:

- ✓ Are the included patients similar to mine?
- ✓ Were clinically important outcomes evaluated?
- ✓ Are the conclusions relevant to my healthcare setting?

EVIDENCE PEARL

A systematic review is only as reliable as the studies it includes and the methods used to conduct the review.

Appraising Clinical Practice Guidelines

Why Appraise Guidelines?

Clinical practice guidelines translate research evidence into **recommendations** for healthcare professionals.

However, not all guidelines are developed with the same methodological rigor.

Key Appraisal Questions:

- ✓ Were all important management options considered?
- ✓ Was an explicit process used to identify and select evidence?
- ✓ Were benefits and harms considered?
- ✓ Were patient values considered?
- ✓ Have recent developments been incorporated?
- ✓ Was the guideline externally reviewed?

Clinical Application

A recommendation should be followed only when:

- The supporting evidence is trustworthy
- The recommendation is applicable to your patient

EVIDENCE PEARL

Do not ask only "What does the guideline recommend?" Also ask: "How was this recommendation developed?"

Choosing the Correct Appraisal Tool

Different study designs require different appraisal tools.

Match the Tool to the Study Design:

Study Design	Appraisal Tool
Randomized Controlled Trial	Therapy Appraisal
Diagnostic Accuracy Study	Diagnostic Test Appraisal
Cohort Study	Prognosis/Harm Appraisal
Case-Control Study	Harm Appraisal
Systematic Review	Systematic Review Appraisal
Clinical Practice Guideline	Guideline Appraisal

Why Does This Matter?

Each study design has different risks of bias and therefore requires different appraisal questions.

EVIDENCE PEARL

The right appraisal tool asks the right questions for the right study design.

CHAPTER TAKEAWAY

- ✓ **Understand** that systematic reviews and guidelines require critical appraisal
 - ✓ **Recognize** that higher levels of evidence are not automatically trustworthy
 - ✓ **Apply** appropriate appraisal tools for different study designs
 - ✓ **Remember** that methodology matters more than study hierarchy
-

Module Summary

Critical Appraisal Framework

Every research paper should be evaluated using three fundamental questions:

1. Is the study valid?

↓

2. What are the results?

↓

3. Can I apply the findings to my patient?

Throughout This Module You Learned To:

- ✓ **Appraise** therapy studies
 - ✓ **Appraise** diagnostic studies
 - ✓ **Appraise** prognosis studies
 - ✓ **Appraise** harm studies
 - ✓ **Appraise** systematic reviews
 - ✓ **Appraise** clinical practice guidelines
 - ✓ **Select** the appropriate appraisal tool for different study designs
-