Critical appraisal for medicine and health sciences

In this post we explore the process of critical appraisal, types of study and the key areas to examine when appraising a research article.

# Introduction

There’s a lot of research out there. [Medline](https://www.librarysearch.manchester.ac.uk/discovery/fulldisplay?vid=44MAN_INST%3AMU_NUI&tab=local&docid=alma9925653194401631&context=L&lang=en) alone provides access to 5000 biomedical journals, with daily updates to the 10 million research articles. Papers may vary in quality, and many will not be relevant to everyday practice.

It’s both impossible and unnecessary to read all the research published in your field. How can you decide which ones you can safely discard?

Understanding the principles of critical appraisal will equip you with the tools to quickly and efficiently identify which articles you should read to inform your practice. Critical appraisals are specific to health science and medical research.

This post explores the process of critically appraising research papers, focusing on these areas:

1. What is critical appraisal?
2. Where do I need to look?
3. Types of study design
4. What do I need to look for?

# What is critical appraisal?

We need to be able to critically appraise research literature in order to be sure that our clinical practice is based upon the best available research evidence.

Critical appraisal has been defined as:

“The process of assessing and interpreting evidence by systematically considering its validity, results and relevance to an individual’s own clinical work.” (Last, 1988)

The majority of papers published in medical, healthcare and dental journals follow the IMRAD format ([Greenhalgh, 1997](https://www.mazums.ac.ir/Dorsapax/Data/Sub_30/File/read%20papers.pdf)).

* Introduction
* Methods
* Results
* Discussion

Most papers also have an abstract at the beginning, which gives an overview of the key elements of each section. In order to decide if an article is worth reading, where would you focus your attention?

## Abstract

The abstract provides a summary of a paper. It can be useful when deciding whether a study is relevant. However, it should not be used to make judgements concerning the validity of the research or results. It is estimated that 18–68% of medical journal abstracts contain omissions or inaccuracies ([Pitkin, 1999](https://pubmed.ncbi.nlm.nih.gov/10188662/)).

## Introduction

The introduction should identify gaps in current knowledge in the area that the paper investigates. It should also outline the aim of the study. As with the abstract, an introduction is useful to establish the relevance of a study, but it will not tell you anything about the validity of the study or what the results are.

## Methods

The methods section outlines how the study was conducted. If you have established the topic of the paper is relevant, the methods section will enable you to identify the validity of the study. This can indicate whether the paper is worth reading.

## Results

The results section will document the findings of the study. This section will be of interest if you have previously identified from the abstract and methods sections that the study is relevant and valid.

## Discussion/Conclusion

The discussion or conclusion section does not always represent the actual findings of the study. For your own research, it is important that you focus on the methods and results sections and think about what the findings mean to you and your practice.

When you have a limited amount of time, it can be very tempting to focus just on the abstracts and results or conclusions of a study. However, to evaluate if a paper would be valuable to read pay attention to the methods section to establish the validity of the study design. Once you are satisfied with the validity of the study, then consider the results and whether it is relevant to your clinical practice.

# Types of study

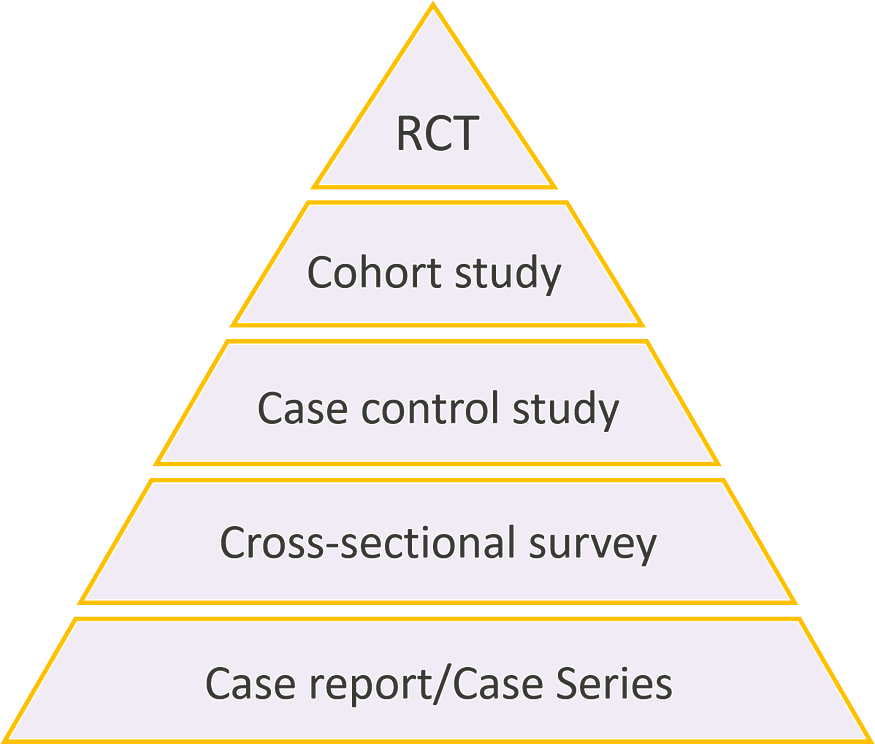
Before you start to appraise a research paper, it is helpful to know what type of study it is. Different study types have their own strengths and weaknesses. Being able to identify the type of study can give you an overall indication of the quality of the research.

There are several critical appraisal checklists available online to help you with appraising distinct types of study. Examples of these include [CASP,](https://casp-uk.net/casp-tools-checklists/) [CEBM](https://www.cebm.net/2014/06/critical-appraisal/), and [AMSTAR](http://amstar.ca/Amstar_Checklist.php). Sometimes a research paper may use more than one method of research

Subject databases such as Medline may include the publication type in the Complete Reference entry for the article. However, sometimes this is not listed, and you will need to consult the research paper to find this.

# Hierarchy of evidence

The hierarchy of evidence can be used as a general indication of the quality of a piece of research. There is no universally accepted hierarchy, and there are many variations of the hierarchical pyramid. The diagram below represents an accepted illustration of the relative strengths of a few of the key types of study.



Hierarchy of evidence.

## Randomised Controlled Trial (RCT)

* **What is it?** A comparison of two or more groups of patients who have been randomly assigned to an experimental or control group. The experimental group is exposed to a treatment or intervention of interest, and the control group are exposed to another treatment or intervention or none at all. The exposure of interest is controlled by the researcher.
* **Example:** How effective is intravenous magnesium for the treatment of acute migraine?
* **What’s it used for?** Evaluating the effectiveness of an intervention.
* **What are its limitations?** They are expensive to conduct. There are ethical issues associated with such clinical trials.

## Cohort study

* **What is it?** Two groups of patients (cohorts) are identified, one of which has received the exposure being studied, one of which has now. Both cohorts are monitored to assess the outcome being studied.
* **Example:** The association between obesity and prostate cancer.
* **What’s it used for?** Measuring the incidence of a disease or condition. Examining the causes of a disease or condition. Establishing the timing and directionality of events.
* **What are its limitations?** The selection of control groups can be difficult. For rare diseases, a large sample size and/or long follow-up is necessary.

## Case control study

* **What is it?** A set of patients with a defining characteristic of interest is selected and compared with a control group without the characteristic.
* **Example:** Increased risk of dementia in people with previous exposure to general anaesthesia.
* **What’s it used for?** Investigating the potential causes of diseases and conditions, particularly rare ones.
* **What are its limitations?** The selection of control groups can be difficult. It is difficult to establish time relationships between exposure to the risk factor and the development of the disease.

## Cross sectional survey

* **What is it?** An observation of a defined population at a single point in time or over a specific time interval.
* **Example:** What is the prevalence of orofacial pain in people living in Manchester?
* **What’s it used for?** Measuring the prevalence of a disease. Examining potential risk factors or causes.
* **What are its limitations?** It cannot establish causality, only association at the most. Surveys can be unreliable due to recall bias of participants.

## Case reports

* **What is it?** A report on a single patient or a series of patients. No control group involved.
* **Example:** A doctor publishes details of a case regarding the birth of two babies with absent or malformed limbs. Both mothers had taken Thalidomide. The report triggers research into the drug.
* **What’s it used for?** Recognition of the new diseases, conditions or outcomes. Formulation of hypotheses.
* **What are its limitations?** It cannot demonstrate a valid statistical association.

# Systematic Reviews

## What type of study is a Systematic Review?

Systematic reviews are usually exhaustive studies which seek to answer a single research question. Thinking about the hierarchy of evidence a Systematic Review would sit above the pyramid as it brings together all reliable data that is synthesised to form an overall consensus or response, to the research question. Evidence may be taken from a variety of sources that sit within the hierarchical pyramid.

It is important to note that Systematic Reviews do have other impingements placed upon them depending on who is conducting them and how long the research period is. For example, an Undergraduate completing a Systematic Review as part of a dissertation project would not be expected to produce a Systematic Review to the same depth and intensity of someone completing one for a PhD or as part of a collaboration of researchers who may take several years to conduct their research and synthesise their findings.

The hierarchical pyramid illustrated is one such example of broad information types. There are other more complex pyramids which may vary slightly from this one.

# What do I need to look for?

There are three key areas to examine when appraising a research article:

## 1. Validity: are the methods robust?

In order to establish whether an article is worth reading or not, we need to examine its validity. This includes a number of factors:

**Study design:** Is the study design appropriate for the research topic or question? Remember:

* Experimental studies are less susceptible to bias than observational studies.
* Prospective studies are less susceptible to bias than retrospective studies.
* Controlled studies are less susceptible to bias than uncontrolled studies.

**Participants:** How have the participants been selected?

**Care provided:** What care was provided to the patients?

**Outcome assessment:** How were the outcomes of the study assessed

**Dropouts:** Were all the participants accounted for?

**Other issues:** How was the sample size calculated? What was the duration of the follow-up?

## 2. Results: are the findings credible?

In order to appraise the results of a study, we need to consider:

* **Completeness:** Are the results presented for all outcomes? Are there any gaps in the results?
* **Reliability:** Do the numbers all add up? Does the data stand alone?
* **Significance:** Are the results statistically significant? If so, are they clinically important?

## 3. Relevance: does it apply to your practice?

Not all valid research with credible results will be relevant to your practice. We need to establish:

* **Applicability:** Can the findings be used more generally? Can you apply them to your own practice?
* **Feasibility:** How feasible is it to implement the findings in your own practice?

# Summary

Your first attempt at critical appraisal may feel slow, but with practice it becomes much quicker and almost automatic. It is important that healthcare workers wanting to engage in evidence-based practice can quickly identify high-quality, clinically relevant research articles to inform their work. By developing critical appraisal skills, you will become better at managing information overload.

# References

Last, J.E. (1988). *A dictionary of epidemiology*. New York: Oxford University Press.

Greenhalgh, T. (1997). ‘How to Read a Paper: Getting Your Bearings (Deciding What the Paper Is about). *The BMJ*, 315 (7102), pp. 243–246.

Pitkin RM, Branagan MA, & Burmeister LF. (1999). ’Accuracy of Data in Abstracts of Published Research Articles’. *JAMA,* 281 (12), pp. 1110–1111.