

Critically Appraising Knowledge for Clinical Decision Making

Grading the strength of a body of evidence should incorporate three domains: quality, quantity and consistency (Agency for Healthcare Research & Quality, 2002)

Quality: The extent to which a study's design, conduct and analysis has minimized selection, measurement and confounding biases

Quantity: the number of studies that have evaluated the question, overall sample size across all studies, magnitude of the treatment effect, strength from causality assessment such as relative risk or odds ratio.

Consistency: whether investigations with both similar and different study designs report similar findings

Questions to Consider

What were the results of the study?

Are the results valid? (validity)

What are the results? Is there clinical and/or statistical significance? (reliability)

Will the results help me in caring for my patients? (applicability)

Validity of the study refers to whether the results of the study were obtained via sound scientific methods. The clinician must be aware of possible sources of bias.

Selection bias: How was the selection of the study participants chosen? Were they representative of general population? Were they randomized?

Measurement bias: Systematic error can occur through using incorrectly calibrated devices resulting in inaccurate readings. A loss of subjects to follow-up may cause what is identified as **study attrition**. The proportion of people who started the study but do not complete the study, for whatever reason. Non-reporting of losses may mask the real reason for observed differences between the experimental group and control group.

Contamination: This may occur when participants in one group are exposed to interventions from the alternate group.

Confounding variables: Those factors that interfere with the relationship between the independent variables (intervention or treatment) and dependent variables (outcome of cause).

Reliability of study findings refers to whether the effects have sufficient influence on practice, clinically and statistically. The **confidence interval (CI)** is used to identify the degree of precision, the degree of random error. The confidence interval gives us the range in which the real answer lies with a given degree of certainty. In general, research papers will report a 95% CI. A 95% CI for the mean indicates having a 95% probability of containing the true mean.

Sample size counts: The *wider* the CI, the less confident we are about the true result. The width of the CI is greatly dependent on sample size. The larger the sample size in a study, the greater the power is to detect a true result and the more confident we can be about the study findings, resulting in a smaller CI.

Statistical significance: P values are the statistical test of the assumption that there is no difference between an experimental intervention and a control. A p value of 0.05 or less is considered a statistically significant result. A p value of 0.05 or less means it is very unlikely that the observed result occurred by chance.

Sensitivity: refers to the proportion of people who have positive test results and who really have the disease.

Specificity: refers to the people who do not have the disease and whose test results are negative.

If a diagnostic test has a **specificity of 1.0** and a **sensitivity of 1.0**, it is said to be a perfectly accurate and valid test.

Why was the study done?

How was the sample size decided?

Are the measurements valid and reliable?

How were the data analyzed?

Were there any untoward events during the conduct of the study?

How do the results fit with previous research in the area? Do the findings compliment or contradict previous work?

What does this research mean for clinical practice?