

CRITICAL READING OF THE MEDICAL LITERATURE
An introduction

by

Yaw Amoateng-Adjepong, MD MPH PhD
Department of Medicine-Pediatrics
Bridgeport Hospital, Yale-New Haven Health

CARDINAL ISSUES

- **Relevance**
 - As assessed in the reader's context
 - Specific study objectives

- **Informational**
 - Adequacy of data for meaningful interpretation

- **Validity**
 - Study type
 - Subject selection
 - Interventions
 - Information gathering/ measurements
 - Extraneous factors
 - Data analysis
 - Report of study's findings

- **Inference** (In the context of the individual study)
 - Chance/Random error
 - Bias
 - Confounding
 - True association (causal or non-causal)

- **Generalization**
 - Extension of study results to the "general" population
 - External consistency

- **Applicability**
 - Practice/Patient-specific applicability

- **Reaction**
 - What to do with the information

RELEVANCE:

- Usually determined from title, objectives and summary
- Reader-specific
- Does it deal with issues of importance in one's practice
- Is it likely to impact practice patterns or challenge beliefs

INFORMATIONAL

- Is the study of reasonable size for meaningful inference?

*Study size:

Is it sufficient to detect a meaningful effect if one truly exists? What is the POWER of the study? Power depends on : Effect size, Study size and the accepted Type 1 error . As a rule, a study must have at least 80% power. (Type 1 error = declaring that a difference exists when it does not. Alpha = probability of committing type 1 error)

- Adequacy of data gathering
- Missing values?

VALIDITY:

- Central to any critical analysis of a paper
- How accurately/truthfully does the study evaluate, assess or measure what it purports to do
- Mere publication by an "expert" or in a premier journal is no guarantee of a study's validity
- Assessment of validity must be thorough. Systematic errors can occur at every stage of the research process, even in the selective publication of research findings.

ASSESSMENT OF VALIDITY

*Objectives/Hypothesis:

Reduce study's goals to clear, verifiable (testable) objectives/hypotheses

*Study Design:

Given the stated objectives, is the chosen study design appropriate and efficient?
Certain types of studies are prone to specific types of systematic errors.

Study Types:

EXPERIMENT: Requires active intervention

Clinical Trial - Patients as study subjects

Field Trial - Healthy people as the study subjects

Community Intervention Trial - intervention assigned to groups of healthy subjects.

NONEXPERIMENTAL STUDIES:

Follow-up (Cohort) Studies ► Can be prospective, retrospective or

ambispective

Case-Control/Case-cohort/Nested case-control studies

Ecologic studies ► unit of observation is group of individuals rather than an individual

Cross-sectional surveys

Case series reports

Hybrid studies

REVIEWS:

Qualitative (narrative) reviews

Quantitative (systematic) reviews / meta analyses

Pooled studies

Practice guidelines

*Subject selection (prevention of selection bias)

-Are the index and comparison subjects similar in their baseline risks for the outcomes of interest?

-Was assignment of index and comparison categories influenced by knowledge of or correlates of the outcomes of interest?

-Exclusion criteria - were these selectively applied?

-Sampling: could the methods used have introduced differences in baseline risk between index and comparison subjects?

- What measures (e.g. Randomization, matching, subject characteristic restrictions) were instituted to minimize differences in baseline risk

*Information gathering (prevention of information bias)

-Exposure/intervention ascertainment

-How accurate is the classification

-Was exposure meaningfully quantified

-Was an objective measure of exposure, e.g. a biologic marker, utilized

-Were the index subjects truly exposed

-Did the intervention group take the study medication as prescribed?

-What is the possibility of contamination of exposure/intervention outside the study's setting?

*Outcome determination

-Were the endpoints hard or soft?

-How accurate was the determination of the presence or absence of the outcome of interest

- Did knowledge of exposure/intervention status or its correlates influence outcome determination?

-Were study subjects and/or researchers blinded to exposure status?

*Extraneous factors (prevention of confounding)

- What are the known correlates of the outcome? Do we have information on their relative distribution in the index and comparison groups?
- What are the known correlates of the exposure? Are these independent predictors of outcome?
- Which factors, other than the exposure/intervention under study, could produce differences in outcome under the given study's settings?

DATA ANALYSIS

- Do the summary measures truly capture the relationship between exposure and the outcome?

COMMON SUMMARY MEASURES

Frequency measures:

- Incidence (incidence density)
- Prevalence
- Risk (cumulative incidence)
- Proportionate

Measures of effect

Absolute Effects

- Rate difference
- Risk difference
- Prevalence difference
- Mean (Survival) difference

Relative Effects

- Rate ratio
- Risk ratio
- Prevalence ratio
- Odds ratio
- Likelihood Ratio (Bayesian factor)
- Proportional Morbidity/Mortality Ratio (PMR)
- Standardized Mortality (Morbidity) Ratio (SMR)

Measures of impact

- Attributable risk (Attributable risk percent, Etiologic fraction)
- Preventive fraction
- Number Needed to Treat (NNT)
- Number Needed to Harm (NNH)

Diagnostic tests

- Sensitivity

Specificity
Predictive values
Likelihood ratios
Receiver operator characteristic (ROC) curves

- *Refutation of selection bias
 - distribution of baseline characteristics
- *Refutation of information bias
 - quantification of exposure and outcome misclassifications
 - assessment of recall bias
- *Refutation of confounding
 - stratification by confounder groups
 - statistical modeling to adjust for confounders
- *Assessment of effect-modification
- *Assessment of biologic gradient (dose-response)
- *Assessment of latency, induction time and duration effects
- *Sensitivity analysis incorporating maximum possible errors in measurement

ASSESSMENT OF RANDOM ERROR

- *Confidence intervals:
 - provide range of estimates consistent with the observed parameter estimate
 - provide evidence on precision
- *P values most widely used.
 - Credibility of a positive result is enhanced by a small p.
 - P-value has major limitations:
 - Depends on strength of the estimate and sample size (and variance)
 - Has no meaning when a null or near null result is obtained. Credibility of a null result is not enhanced by a large, uninterpretable p.
 - The 0.05 cut-off point is arbitrary
 - Interpretations of p by many assume chance and causality as the only two possibilities
 - Chance can never be excluded even in the absence of a small value. P cannot be zero.
- *P-value function
- *Bayesian factor

INFERENCE - (In the context of a single study)

- *Must address the specified hypotheses/objectives
- *Findings interpretable in terms of:
 - Random error (chance)
 - Bias (selection, information bias)
 - Confounding (extraneous factors)
 - True Association (causal or non-causal)
- *Interpretations are not mutually exclusive
- *Fraud is not usually considered though possible
- *Failure to discount chance does not mean the absence of evidence for association

CRITERIA FOR CAUSATION (in the context of a single study)

- *Minimal bias
 - *Minimal confounding
 - *Minimal random error
 - *Strength of association
 - *Internal consistency
 - *dose-response
 - *uniformity of results
 - *Temporality
 - *Biologic plausibility
- “Causal inference cannot attain the certainty of logical deductions”

*Consider Significance of findings

Clinical Significance ► importance of findings for changing current clinical practice.
Depends on the effect size (the magnitude of clinical response)

Statistical Significance ► relevant only in excluding chance as an explanation for findings

Biologic Significance ► findings help to clarify mechanism of action

GENERALIZATION -Extension of study results beyond the study population. Consider all available scientific evidence on the issue. A study's generalizability may be limited by the narrow definition of the population base on whom information is available.

Useful criteria include:

- *External consistency

- replication (duplication and/or similarity of results among studies)
- strength of association from the varied studies
- *Coherence - congruence of the current evidence with all known relevant biologic information
- *Response to manipulation - not always applicable

For non-human studies, consider

- *Dose equivalence
- *Difference in pathways
- *Products of metabolism

APPLICABILITY OF RESULTS

- *Similarity of patient characteristics. Would your patient have qualified as a study participant?
- *Modifying factors
- *patient compliance
- *provider compliance
- *feasibility of necessary follow-up
- *co-morbid conditions
- *Differences in metabolism
- *possible drug interactions
- *cost versus risk of adverse outcome

REACTION:

What to do with the study's findings. Consider the quality of evidence and the spectrum of credibility. Always maintain an open mind.