

Frederick L. Brancati, MD, MHS

Professor of Medicine & Epidemiology; Director, Division of General Internal Medicine

Welch Center for Prevention, Epidemiology, and Clinical Research

Study Design

Each design has a distinct pattern of advantages and disadvantages that circumscribe the questions it might address and the conclusions it might support. All studies below are “observational” except for randomized, controlled trials. Only trials can prove causality; observational studies can only suggest causal relationships.

Original Research

- ! Case Report Initial report of new therapy or etiologic association
- ! Case Series Possibly compelling report of new therapy, if natural history well established
- ! Ecologic Study Only group data; suggests possible etiologic associations; ecologic fallacy
- ! Cross-Sectional Study “Snap-shot” study; no intervention, no follow-up data
 - Prevalence e.g. How common is osteoarthritis of the hip in Americans aged 65+?
 - Diagnostic tests e.g. Which is better to dx DVT: fibrinogen scan or ultrasonogram?
 - Etiology e.g. Is serologic evidence of chlamydia exposure associated with CVD?
 - Caveats Reverse causality; Survivor bias; Cannot address prognosis or treatment
- ! Case-Control Study Participants selected based on disease present or absent; Look back at exposures
 - Etiology e.g. Is history of head injury more common in Alzheimers pts than controls?
 - Outcomes Research e.g. Is history of hemocult screening less common in colon CA pts vs controls?
 - Caveats Selection bias; Recall bias; Unable to address prevalence, incidence, or prognosis
- ! Prospective Cohort Study Cohort assembled and studied; then followed over time for changes and events
 - Incidence e.g. Which states suffer the highest incidence of respiratory tract cancers?
 - Prognosis e.g. What factors predict sudden death in pts who present to ER with syncope?
 - Etiology e.g. Does dietary Mg intake independently predict the development of diabetes?
 - Outcomes Research e.g. Does peripheral angioplasty appear to reduce the risk of amputation?
 - Caveats Long, costly; Residual confounding; Confounding by indication
- ! Rand Controlled Trial Cohort assembled, then randomly assigned to intervention vs control
 - Therapy e.g. Does enalapril reduce mortality in persons with CHF?
 - Prevention e.g. Does mammography + breast surgery reduce breast CA mortality ?
 - Caveats Generalizability; Often under-powered; Often rely on intermediate effects

Data Synthesis

- ! Qualitative Review No explicit methods; must trust author’s judgement
- ! Meta-Analysis Explicit methods, e.g. paper selection, assessment of rigor, pooling
 - Clinical events e.g. Does aspirin reduce the risk of first MI?
 - Physiologic traits e.g. Do ACEI’s reduce the progression of albuminuria in diabetic adults?
 - Caveats Depends on homogeneity and study selection
- ! Decision Analysis Create tree; assign utilities & probabilities; calc expected value of decision

Mortality	e.g. Does lipid rx reduce mortality in elders without history of MI
QALY	e.g. Does prostate CA screening improve duration and quality of life?
Cost-Benefit	e.g. Which echo leads to higher benefit/cost ratio: trans-esoph or trans-thoracic?
Caveats	Many assumptions; costs and utilities hard to measure

! Practice Guidelines Widely variable quality; strongest based on big trial or meta or decision analysis

Sources of Error

! Bias (Systematic error)	Threatens internal validity; Occurs in design, conduct, analysis, inference
Selection Bias	e.g. Controls in panc CA-coffee study less likely to drink coffee than average
Information Bias	e.g. Mothers ruminate over ASA use in study of Reye's syndrome
	e.g. Silent ischemia prompts greater ECG vigilance in PCS of peri-op MI risk
Analytic Bias	e.g. Non-compliers dropped from analysis in trial of resins to prevent MI
Remedies	Pop-based controls, Standardization; Blinding; Intention-to-treat analysis
! Chance (Random error)	Reduces statistical power; precision; and statistical significance
Variability	e.g. BP differs day-to-day in same individual
Uncommon Events	e.g. ESRD occurs in only 14 of 10,000 adults over 10 yr interval
Remedies	More precise measurements; repeated measures; Larger N; Longer follow-up
! Confounding	Real association, wrong inference; e.g. grey hair 6 CAD (confounded by age) Residual
Unsuspected	e.g. Alcohol 6 Lung CA if higher smoking rate in drinkers not accounted for
Remedies	Matching; Collect data on confounders and conduct multivariate analysis

From Journal to Bedside

! All Study Designs	
Internal Validity	The data support the conclusions in study sample
External Validity	Conclusions are generalizable to other populations (or your patient)
Statistical Significance	It's unlikely that the results occurred by chance ($p < 0.05$)
Clinical Significance	The results are compelling enough to influence your practice
! Rx or Prevention	
Efficacy	Whether the intervention works under ideal circumstances (controlled trials)
Effectiveness	Whether the intervention works in real-world circumstances (outcomes research)
Intermediate Effects	Disease-related traits (e.g. HbA1c, BP, FEV1, Ejection fraction)
Outcomes / Events	e.g. Renal failure, Stroke, Hospitalization for COPD, Death
Rel Risk Reduction	$(\text{Incidence in treated}) / (\text{Incidence in controls})$
Abs Risk Reduction	$(\text{Incidence in controls}) - (\text{Incidence in treated})$
N Needed to Treat	$1 / (\text{Absolute Risk Reduction})$
! Diagnostic Tests	
Sensitivity	If he has it, how likely is test to yield positive result?
Specificity	If he doesn't have it, how likely is test to yield negative result ?
Pre-test Probability	How likely do I think it is, even before I order the test?
Post-test Probability	How likely is it in light the test results?
Pos Predictive Value	Given that test is positive, how likely is it? (Depends strongly on pre-test prob)
Odds	$1 / (1 - \text{probability})$
Likelihood Ratio	Test characteristic, strong test $LR \gg 1$ or $\ll 1$; useless test $LR \approx 1$
Post-test Odds	Pre-test Odds x LR; to convert to post-test prob, $1 / (\text{odds} + 1)$

Glossary

Blinding	Keeping staff unaware of exposures or treatments to improve even-handedness
Cohort	Well-defined group followed over time for changes, incident disease, or death
Ecologic Fallacy	When only group data are available, conclusions about individuals may be wrong
Incidence	New cases of disease in a group at risk / [N in group x defined time interval]
Prevalence	Existing cases of disease in a group / N in group
Multivariate Analysis	Use of regression (e.g. logistic) to adjust simultaneously for several confounders
Odds	Measure of risk related to probability, but easier to manipulate mathematically
Recall Bias	Tendency of rumination to change reporting after a clinical event has occurred
Reverse Causality	e.g. In cross-sectional study, donuts appear “protective” against diabetes
Power	Ability of a study to detect a given association; must know when study is negative
Standardization	Uniform approach to measurement, definition, etc in order to reduce bias
Survivor Bias	In octogenarians, smoking weakly assoc w/ COPD (since susceptibles already died)
Tree	Diagram of decision and its many consequences, ending in death, costs, or utilities
Utility	Value placed on given health state by patient, doctor, general population, etc

Contingency Tables

Analytic approach adaptable for almost any study design (aka 2 x 2 table)

	Disease		Totals
	Yes	No	
Exposed / Treated / Positive	a	c	a + c
Unexposed / Control / Negative	b	d	b + d
Totals	a + b	c + d	a+b+c+d

In Cross-Sectional study

$$\text{Prevalence} = (a + b) / (a + b + c + d)$$

$$\text{Odds Ratio} = (a / b) / (c / d) = ad / bc$$

for assoc of exp with disease

In Case-Control Study

Prevalence cannot be calculated

Odds Ratio calculated as above

In Prospective Cohort Study

$$\text{Incidence in exp} = a / [(a + c) \times t]$$

$$\text{Incidence in unexp} = b / [(b + d) \times t]$$

$$\text{Rel Risk} = \text{Inc in exp} / \text{Inc in unexposed}$$

In Diagnostic Test Study

$$\text{Sensitivity} = a / (a + b)$$

$$\text{Specificity} = d / (c + d)$$

$$\text{Pre-test Prob} = (a + b) / (a + b + c + d)$$

$$\text{Pre-test Odds} = (a + b) / (c + d)$$

$$\text{Pos Pred Value} = a / (a + c)$$

In Randomized Controlled Trial

$$\text{Rel Risk Red} = \text{Inc in rx'd} / \text{Inc in nrx'd}$$

$$\text{Abs Risk Red} = \text{Inc in rx'd} - \text{Inc in unrx'd}$$

Caveat: Watch for different table lay-outs,
i.e. rows and/or columns switched