

CORE CURRICULUM IN NEPHROLOGY

Study Designs in Patient-Oriented Research

Chirag R. Parikh, MD, PhD, and John Concato, MD, MS, MPH

A variety of study designs are used in patient-oriented research. Selected designs are more suitable in certain situations, and it therefore is important for researchers to understand the advantages and disadvantages of various study designs to apply them appropriately. The characteristics of a disease (eg, prevalence, acute versus chronic course) and the research question being addressed influence the corresponding study design (architecture) for that study. [Figure 1](#) gives the overview of research architecture, and the subsequent text gives salient features, strengths, and limitations of each study design. [The Appendix](#) provides a glossary of commonly used terms.

DESCRIPTIVE STUDIES

Case Report/Case Series

- Case report: a brief description of a single case that an observer thinks should be brought to colleagues' attention
- Case series: several case reports of similar observations, procedures, etc, that can be grouped together
- Simple to perform; report usually can be written up and published rapidly
- Often first form of reporting for new diseases or rare complications
- Very limited in discerning cause-effect relationship or comparison of treatment effects (unless procedure is dramatic [eg, first successful dialysis])

From the Section of Nephrology, Department of Medicine and Clinical Epidemiology Research Center, Yale University School of Medicine and Veterans Affairs Medical Center, New Haven and West Haven, CT.

Received August 1, 2005; accepted in revised form September 13, 2005.

Originally published online as doi:10.1053/j.ajkd.2005.09.035 on January 4, 2006.

Address reprint requests to Chirag R. Parikh, MD, PhD, Section of Nephrology, Yale University and VAMC, 950 Campbell Ave, Mail Code 151B, Bldg 35 A, Rm 219, West Haven, CT 06516. E-mail: chirag.parikh@yale.edu

© 2006 by the National Kidney Foundation, Inc.

0272-6386/06/4702-0019\$32.00/0

doi:10.1053/j.ajkd.2005.09.035

Incidence/Prevalence Studies

- Determines magnitude of disease or a disease characteristic in population
- Very important for planning resource utilization
- Usually cross-sectional (prevalence) or longitudinal (incidence) design

Examples of Descriptive Studies

- Merrill JP, Murray JE, Harrison JH, Guild WR: Successful homotransplantation of the human kidney between identical twins. *JAMA* 160:277-282, 1956
- Hume DM, Merrill JP, Miller BF, Thorn GW: Experiences with renal homotransplantation in the human: Report of nine cases. *J Clin Invest* 34:327-382, 1955

ANALYTIC STUDIES

Analytic studies are a more commonly encountered category of studies, involving comparisons between 2 or more groups. They are based on a research question and are etiologic, diagnostic, prognostic, therapeutic, and so on. Based on research architecture, the studies can be observational or experimental.

Observational Studies

In observational studies, exposure is not determined by the investigator.

Cross-sectional studies

- No longitudinal component ([Fig 2](#))
- Data on exposure and outcome assessed at same time
- Strength: no waiting (fast and inexpensive)
- Weakness: limited inference between exposure and outcome
- Caution: potential biases (eg, response/participation bias [sicker patients are more or less likely to participate])

Examples of cross-sectional studies

- Kramer HJ, Nguyen QD, Curhan G, et al: Renal insufficiency in the absence of albumin-

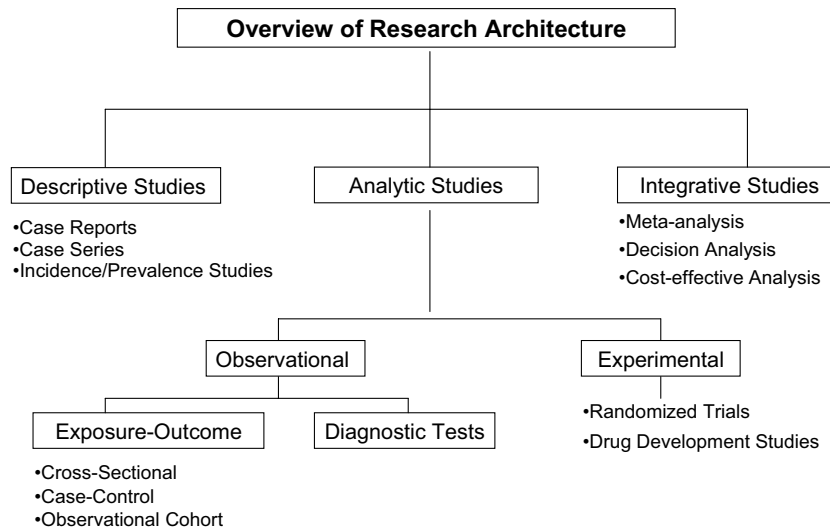


Fig 1. Overview of research architecture.

uria and retinopathy among adults with type 2 diabetes mellitus. JAMA 289:3273-3277, 2003

- Nash D, Magder L, Lustberg M, et al: Blood lead, blood pressure, and hypertension in perimenopausal and postmenopausal women. JAMA 289:1523-1532, 2003

Cohort studies

- Follows up groups of participants (eg, exposed and nonexposed) over time (Fig 3)
- Describes natural history, establishes temporal sequence
- Estimates of incidence and relative risk can be obtained
- Two common variations, prospective and retrospective studies, depending on whether

exposure and outcome occurred concurrent with research or prior to research, respectively:

Prospective cohort design.

- Useful strategy to assess cause-effect relationship
- Important design for predictors that can otherwise have recall bias (eg, diet, alcohol)
- Requires large sample size and long follow-up periods
- Allows calculation of relative risk for given exposure and outcome
- Limitations: inefficient method for studying rare diseases; confounding factors can threaten validity of findings

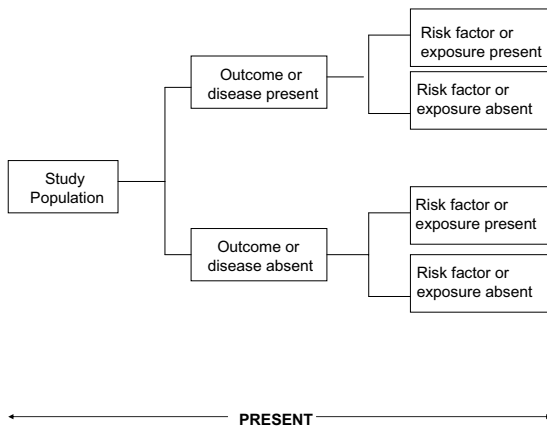


Fig 2. Cross-sectional study.

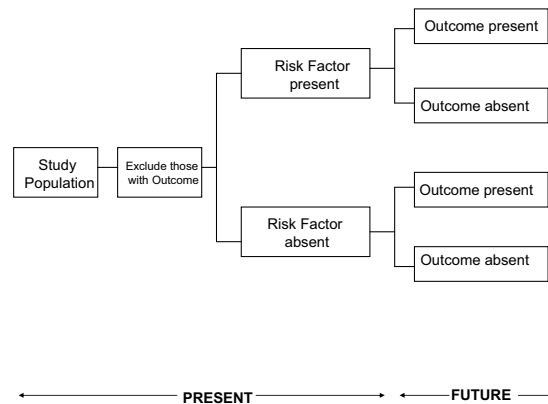


Fig 3. Cohort study.

- Caution: preventing losses during follow-up is important during conduct of study for accurate results; other biases (surveillance bias [more procedures in group with exposure])

Retrospective cohort design.

- Much less time-consuming and costly compared with prospective cohort studies
- Limited control over sampling of cohort and quality of predictor variables (collected in past)

Propensity score methodology.

- Improves causal inference in observational studies when compared groups are different at baseline with respect to intervention (exposure)
- Propensity score is model-based predicted probability of receiving intervention (exposure)
- Controls for confounding and minimizes loss of degrees of freedom in statistical analyses
- Outcomes in intervention and control group are “weighted” across values of propensity scores

Examples of cohort studies

- Go AS, Chertow GM, Fan D, et al: Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. *N Engl J Med* 351:1296-1305, 2004
- Ojo AO, Held PJ, Port FK, et al: Chronic renal failure after transplantation of a nonrenal organ. *N Engl J Med* 349:931-940, 2003
- Abbott KC, Trespalacios FC, Agodoa LY, Taylor AJ, Bakris GL: Beta-blocker use in long-term dialysis patients: Association with hospitalized heart failure and mortality. *Arch Intern Med* 164:2465-2471, 2004
- Kayler LK, Rasmussen CS, Dykstra DM, et al: Gender imbalance and outcomes in living donor renal transplantation in the United States. *Am J Transplant* 3:452-458, 2003

Case-control studies

- See Fig 4
- High yield of information from relatively few subjects

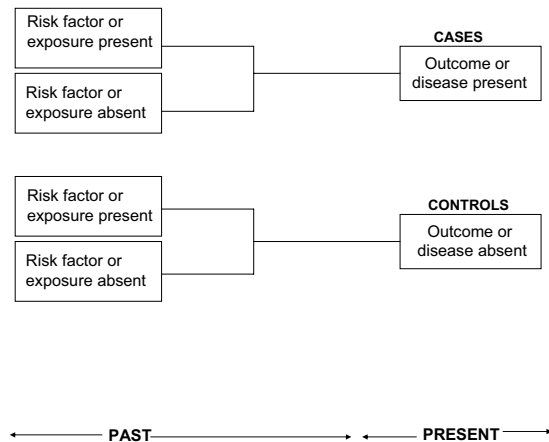


Fig 4. Case-control study.

- Efficient design for rare outcomes
- Can be used to generate or test hypothesis
- Estimates odds ratio; estimates strength of association between predictor and disease (outcome); odds ratio is good approximation of relative risk if outcome is rare
- Limitations:
 - Cannot evaluate incidence/prevalence
 - Only 1 outcome usually studied
 - Limited control over quality of predictor variables
 - Increased susceptibility to bias: sampling bias (selection of controls from unrelated population), recall bias (smoking or alcohol intake history)
- Nested case-control design:
 - Case-control study embedded within cohort study
 - Particularly useful design for predictors (eg, biomarkers) that are expensive to measure in whole cohort
 - Preserves advantages of cohort studies (ie, cases and controls are from same population and predictors are measured before outcomes)
 - For biomarker studies, samples need to be stored until outcomes occur for diagnostic studies

Examples of case-control studies

- Keller G, Zimmer G, Mall G, et al: Nephron number in patients with primary hypertension. *N Engl J Med* 348:101-108, 2003
- Alfrey AC, LeGendre GR, Kaehny WD: The dialysis encephalopathy syndrome. Pos-

- Follow-up visit description and schedule
- Ascertainment of response variables: training, data collection, quality control
- Data analysis: interim monitoring, final analysis
- Termination policy
- Organization:
 - Participating investigators: statistical unit or data-coordinating center, laboratories and other special units, clinical centers
 - Study administration: steering committees and subcommittees, data monitoring and safety committee, funding organization
- Limitations:
 - Costly in time and money
 - Some research questions not suitable for experimental design (eg, etiology, adverse effects)
 - Research interventions may not be feasible in clinical practice
 - Can involve restricted scope and narrow study question
- Modifications of randomized trials:
 - Factorial design: 2 separate questions in single cohort of patients
 - Randomization of matched pairs
 - Randomization of groups or clusters
 - Crossover design: intervention followed by placebo in 1 group and converse in another group:
 - Limitation: carryover effect
 - Adequate “washout” periods needed
 - Trials for Food and Drug Administration approval of new therapies:
 - Phase I: unblinded, uncontrolled studies in few volunteers to test safety
 - Phase II: relatively small, randomized, blinded trial to test tolerability and different intensity or dose of intervention on surrogate outcomes
 - Phase III: relatively large, randomized, controlled, blinded trials to test effect of therapy on clinical outcomes
 - Phase IV: large trials or observational studies, conducted after intervention has Food and Drug Administration approval, to assess rate of serious side

effects and evaluate additional therapeutic uses

Examples of Experimental Studies

- Lewis EJ, Hunsicker LG, Clarke WR, et al: Renoprotective effect of the angiotensin-receptor antagonist irbesartan in patients with nephropathy due to type 2 diabetes. *N Engl J Med* 345:851-860, 2001
- Brenner BM, Cooper ME, de Zeeuw D, et al: Effects of losartan on renal and cardiovascular outcomes in patients with type 2 diabetes and nephropathy. *N Engl J Med* 345:861-869, 2001

INTEGRATIVE STUDIES

Meta-Analysis

- Systematic review that uses statistical techniques to quantitatively combine and summarize results of previous research
- Rationale:
 - Obtain more precise estimates of intervention effect
 - Enhance statistical power to observe small, but clinically important, intervention effects
 - Opportunity to perform important subgroup analysis
 - Evaluate the generalizability of results across trials and populations
 - Determine whether opportunity exists to conduct new study (loss of equipoise?)
- Performing a meta-analysis:
 - Written research protocol with well-defined question
 - Methods for identifying eligible studies:
 - MEDLINE, PubMed, Web of Science, EMBASE
 - Dissertation libraries (*Index Medicus*)
 - Conference proceedings and abstracts
 - Communication with experts in the area
 - Methods for abstracting data
 - Statistical methods:
 - Summary effect estimate and confidence interval
 - Tests for evaluating heterogeneity and potential publication bias
 - Planned subgroup and sensitivity analysis

- Analysis involves fixed-effects model or random-effects model (preferred)
- Limitations:
 - Publication bias: negative studies infrequently published; abstracts, small studies, and theses difficult to find
 - Language bias: non-English literature difficult to obtain and translate
 - Heterogeneity across different studies important to control:
 - Meta-regression methods
 - Fixed-effects model with covariates
 - Random-effects model with covariates
 - Control rate model

Examples of Meta-Analyses

- Turnbull F: Effects of different blood-pressure-lowering regimens on major cardiovascular events: Results of prospectively-designed overviews of randomised trials. *Lancet* 362:1527-1535, 2003
- Jafar TH, Stark PC, Schmid CH, et al: Progression of chronic kidney disease: The role of blood pressure control, proteinuria, and angiotensin-converting enzyme inhibition: A patient-level meta-analysis. *Ann Intern Med* 139:244-252, 2003
- Garg AX, Suri RS, Barrowman N, et al: Long-term renal prognosis of diarrhea associated hemolytic uremic syndrome: A systematic review, meta-analysis, and meta-regression. *JAMA* 290:1360-1370, 2003

Decision Analysis

- Method for rational decision making that incorporates best available medical, diagnostic, and economic evidence
- Particularly suited for clinical situations that are highly complex; considers patient preferences, physician decisions, etc
- Performing a decision analysis:
 - Formulate question
 - Structure decisions and build decision tree:
 - Decision tree is composed of decision nodes, chance nodes, and terminal states (outcomes)
 - Fill in data (probabilities and outcomes):

- Probabilities obtained from clinical studies, clinical databases, calculated guesses, or expert opinion
- Determine value of each competing strategy
- Perform sensitivity analysis:
 - Sensitivity analysis tests stability of results of decision tree by systematically varying values of particular probabilities and outcomes incorporated within tree
- Advanced methods of decision analysis:
 - Time preferences: involves discounting utility of health in future compared with present
 - Markov process
 - Monte-Carlo simulation
- Cost-effectiveness analysis uses decision analysis concepts for selecting among competing strategies when resources are limited
- Limitation:
 - Not practical for many clinical situations
 - Paucity of data can make study challenging
 - Preferences may be difficult to elicit, may differ from real life situation, or may change over time

Examples of Decision Analyses

- Roberts SD, Maxwell DR, Gross TL: Cost-effective care of end-stage renal disease: A billion dollar question. *Ann Intern Med* 92:243-248, 1980
- Boulware LE, Jaar BG, Tarver-Carr ME, et al: Screening for proteinuria in US adults: A cost-effectiveness analysis. *JAMA* 290:3101-3114, 2003

DESIGNING STUDIES FOR MEDICAL TESTS

- Medical tests are integral to clinical practice and are performed to screen for a risk factor, diagnose a disease, or estimate prognosis
- Most designs for medical tests resemble observational study designs discussed; clinical trials rarely used
- Diagnostic test seeks to determine usefulness in clinical practice rather than determining “causality”
- Analysis focused on confidence intervals, sensitivity, specificity, and test performance, as opposed to statistical significance

(*P* value), which plays major role in other study designs

- Usefulness of diagnostic test is dependent on series of pertinent considerations:
 - Reproducibility:
 - Studies of intra- and interobserver variability
 - Studies of intra- and interlaboratory variability
 - Test performance:
 - Sensitivity
 - Specificity
 - Positive and negative predictive values
 - Receiver operating characteristic curves
 - Likelihood ratios
 - Feasibility:
 - Studies involving cost of test
 - Proportion willing to undergo test
 - Proportion experiencing side effects
 - Impact on clinical decisions and clinical outcomes:
 - Difficult to perform
 - Dependent on availability of treatment or intervention after results of test are obtained
 - Proportion of tests leading to changes in clinical decision making
 - Studies estimating risk ratios, odds ratios, number needed to treat for which the predictor variable is receiving a test
- Caution with:
 - Spectrum bias: test works well only in severe cases
 - Observer bias: can be eliminated by blinding
 - Institution-specific results
 - Interpretation and analysis of borderline and uninterpretable results

Examples of Studies for Medical Tests

- Maisel AS, Krishnaswamy P, Nowak RM, et al: Rapid measurement of B-type natriuretic peptide in the emergency diagnosis of heart failure. *N Engl J Med* 347:161-167, 2002
- Levine RJ, Maynard SE, Qian C, et al: Circulating angiogenic factors and the risk of preeclampsia. *N Engl J Med* 350:672-683, 2004
- Parikh CR, Jani A, Melnikov VY, Faubel S, Edelstein CL: Urinary interleukin-18 is a marker of human acute tubular necrosis. *Am J Kidney Dis* 43:405-414, 2004

ADDITIONAL READING

1. Feinstein AR: *Multivariable Analysis: An Introduction*. New Haven, CT, Yale University Press, 1996
2. Flather M, Aston H, Stables R: *Handbook of Clinical Trials*. London, ReMedica, 2001
3. Friedland DJ: *Evidence-Based Medicine: A Framework for Clinical Practice*. Stamford, CT, Appleton & Lange, 1998
4. Friedman LM, Furberg C, DeMets DL: *Fundamentals of Clinical Trials* (ed 3). New York, Springer, 1998
5. Gordis L: *Epidemiology* (ed 2). Philadelphia, PA, Saunders, 2000
6. Hulley SB: *Designing Clinical Research: An Epidemiologic Approach* (ed 2). Philadelphia, PA, Lippincott Williams & Wilkins, 2001
7. Riegelman RK: *Studying a Study and Testing a Test: How to Read the Medical Evidence* (ed 4). Philadelphia, PA, Lippincott Williams & Wilkins, 2000
8. Rothman KJ, Greenland S: *Modern Epidemiology* (ed 2). Philadelphia, PA, Lippincott-Raven, 1998
9. Concato J, Shah N, Horwitz RI: Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med* 342:1887-1892, 2000
10. Rubin DB: Estimating causal effects from large data sets using propensity scores. *Ann Intern Med* 127:757-763, 1997
11. Concato J: Overview of research design in epidemiology. *J Law Policy* XII:489-507, 2004

Appendix. Glossary of Commonly Used Terms in Patient-Oriented Research

- Absolute risk reduction (ARR):** Mathematical difference in event rates for 2 groups, usually treatment and control.
- Alpha (type I) error:** Error in hypothesis testing when a statistically significant association is found, but no “true” association exists (ie, rejecting the null hypothesis when it is true). The alpha error level is the threshold of statistical significance established by the researcher ($P < 0.05$ by convention).
- Beta (type II) error:** Error in hypothesis testing when no statistically significant association is found, but a “true” association exists (ie, rejecting an alternative hypothesis when it is true). The beta error level is usually set at 0.2 or less.
- Bias:** Systematic error in the design or conduct of a study, which threatens the validity of a study.
- Blinding:** Element of study design in which patients and/or investigators do not know who is in the treatment group and who is in the control group; the term *masking* is often used.
- Confidence interval (CI):** Describes the variability in a point estimate (relative risk, odds ratio, etc); usually reported as a 95% CI (ie, the range of values within which a 95% probability exists for true value).
- Confounding:** A variable having independent associations with both the dependent and independent variables, potentially distorting their relationship.
- Controlling for:** Term used to describe when confounding variables are adjusted in the design or analysis of a study to minimize confounding.
- Cost-effectiveness analysis (CEA):** A form of economic-efficiency analysis in which costs are valued in monetary terms and health benefits are valued in natural units; CEA is incremental, comparing some new health care technology or strategy of interest with a relevant alternative.
- Dependent variable:** Outcome or response variable.
- Distribution:** Values and frequency of a variable (Gaussian, binomial, skewed).
- Effect size:** The magnitude of a difference considered clinically meaningful. Used in power analysis to determine the required sample size.
- Effectiveness:** A measure of the benefit resulting from an intervention for a given health problem under typical conditions of use; this form of evaluation considers both the efficacy of an intervention and its acceptance by those to whom it is offered, providing an answer to the following question: Does the practice do more good than harm to people to whom it is offered?
- Efficacy:** A measure of the benefit resulting from an intervention for a given health problem under ideal conditions of use; it answers the question: Does the practice do more good than harm to people who comply fully with the recommendations?
- Hypothesis:** A formal statement, with statistical implications, that will be accepted or rejected based on the evidence (data collected) in a study.
- Incidence:** Proportion of new cases of a specific condition in the population at risk during a specified time.
- Independent events:** Events whose occurrence has no effect on the probability of each other.
- Independent variable:** Variable associated with the outcome of interest that contributes information about the outcome in addition to that provided by other variables considered simultaneously.
- Intention-to-treat analysis:** Method of analysis in randomized clinical trials in which all patients randomly assigned to a treatment group are analyzed in that treatment group, whether or not they received that treatment or completed the study.
- Interaction:** Relationship between 2 independent variables, such that the effect of 1 variable on the outcome depends on the “level” of the other variable.
- Likelihood ratio (LR):** Likelihood that a given test result would be expected in a patient with a condition compared to a patient without the condition. Ratio of true-positive rate to false-positive rate.
- Matching:** Process of making 2 groups homogeneous regarding possible confounding factors.
- Meta-analysis:** An evidence-based systematic review that uses quantitative methods to combine the results of several independent studies to produce summary statistics.
- Multiple comparisons:** Pairwise group comparisons involving more than 1 P value.
- Negative predictive value (NPV):** Probability of not having the disease given a negative diagnostic test; requires an estimate of prevalence.
- Null hypothesis:** Default statistical hypothesis assuming no difference between groups; a “straw-man” statement that the data will (hopefully) refute.
- Number needed to treat (NNT):** Number of persons who must be treated for a given period to achieve an event (treatment) or to prevent an event (prophylaxis). The NNT is the reciprocal of the absolute risk reduction.
- Odds:** Probability that event will occur divided by probability that event will not occur.
- Odds ratio (OR):** Ratio of the odds of having condition/outcome in experimental group to the odds of having the condition/outcome in the control group; an estimate of relative risk obtained in case-control studies.
- One-tailed test:** Test in which the alternative hypothesis specifies a deviation from the null hypothesis in 1 direction only; eg, treatment can be better (only) than placebo.
- Placebo:** Inactive substance used to reduce bias by simulating the treatment under investigation.

(Continued)

Appendix (Cont'd). Glossary of Commonly Used Terms in Patient-Oriented Research

Positive predictive value (PPV): Probability of having the disease given a positive diagnostic test; requires an estimate of prevalence.

Power: Probability of finding a significant association when one truly exists (1, probability of type II (β) error); by convention, power of 80% or greater is considered sufficient.

Prevalence: Proportion of individuals (in a cross-sectional assessment) with a disease or characteristic in the study population of interest.

Probability: A number, between 0 and 1, indicating how likely an event is to occur.

P value: Probability of type I (α) error. If the *P* value is small, then it is unlikely that the results observed are due to chance.

Random sample: A sample of subjects from the population such that each has equal chance of being selected.

Receiver operating characteristic (ROC) curve: Graph showing the test's performance as the relationship between the true-positive rate and the false-positive rate.

Regression: Statistical technique for determining the relationship among a set of variables.

Relative risk: Proportionate ratio of event rates (regarding therapy, prophylaxis, etc) in the treatment group relative to that in the control group.

Relative risk reduction or increase: Increase in events with treatment compared with control (treatment) or reduction in events with treatment compared with control (prophylaxis); this number is often expressed as a percentage.

Sample: Subset of the population.

Selection bias: Systematic error in sampling the population.

Sensitivity: Proportion of patients who have the outcome that are "test positive."

Sensitivity analysis: Any test of the stability of the conclusions of a health care evaluation over a range of probability estimates, value judgments, and assumptions about the structure of the decisions to be made; this involves the repeated evaluation of a decision model in which 1 or more of the parameters of interest are varied.

Specificity: Proportion of patients without the outcome who are "test negative."

Two-tailed test: Test in which the alternative hypothesis specifies a deviation from the null hypothesis in 2 directions (eg, treatment can be better or worse than placebo).

Validity: Extent to which a questionnaire, instrument, etc, accurately measures what it is intended to measure; or the extent to which a study accurately evaluates what it is intended to evaluate.

NOTE. The definitions have been adapted from commonly used epidemiology textbooks.