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])

**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**

**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-
Overview of Research Architecture

Descriptive Studies
- Case Reports
- Case Series
- Incidence/Prevalence Studies

Analytic Studies

Integrative Studies
- Meta-analysis
- Decision Analysis
- Cost-effective Analysis

Observational
- Exposure-Outcome
- Cross-Sectional
- Case-Control
- Observational Cohort

Experimental
- Diagnostic Tests
- Randomized Trials
- Drug Development Studies

Fig 1. Overview of research architecture.

Cohort studies
- Follows up groups of participants (e.g., 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 (e.g., 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

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

Examples of case-control studies
- Alfrey AC, LeGendre GR, Kaehny WD: The dialysis encephalopathy syndrome. Pos-
Experimental Studies (Randomized Controlled Trials)

- A properly planned and executed randomized clinical trial is a powerful experimental technique for assessing effectiveness of an intervention.
- Randomized, double-blind, placebo-controlled trial (Fig 5):
  - Strongest evidence for cause-effect association
  - Best possible design for some questions
  - Randomization balances predictor variables across treatment groups (protecting against confounders):
    - Block randomization: ensures that participant numbers are equally distributed
    - Stratified randomization: promotes balance of a particular predictor of outcome
    - Other: randomization by site in multisite trial

- Blinding helps avoid bias during data collection and assessment:
  - Unblinded: investigator and participant know about intervention type
  - Single-blind studies: only investigator or patient is aware
  - Double blind: investigator and subject both are unaware

- Components of good clinical trial study protocol:
  - Objectives:
    - Primary aim or question: including predictor and response variables
    - Secondary aim or question: including predictor and response variables
    - Subgroup hypotheses, if any
    - Adverse effects
  - Design of study:
    - Study population: inclusion criteria, exclusion criteria
    - Sample size assumptions and estimates (including statistical power)
    - Enrollment of participants: informed consent, assessment of eligibility, baseline examination, allocation of intervention (method of randomization)
    - Intervention: description and schedule, measures of adherence
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 (e.g., 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

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

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

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


ADDITIONAL READING

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.