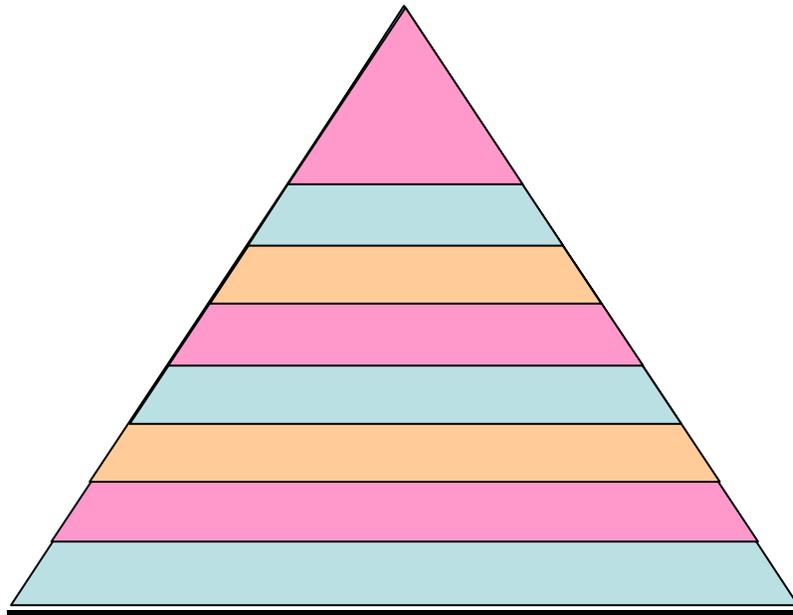


Chapter 5

Appraising the Literature Overview of Study Designs



Barbara M. Sullivan, PhD
Department of Research, NUHS

Jerrilyn A. Cambron, PhD, DC
Department of Research, NUHS

Overview of Clinical Research Study Designs

Randomized, controlled trials (RCT)

- Experimental
- Considered the “Gold Standard” for therapy studies
 - Researcher manipulates the intervention or exposure (independent variable) and records effect on outcome of interest (dependent variable)
- Participants are randomly allocated into intervention (treatment) and control (comparison, placebo) groups
 - Randomization (if done) method is key to RCT; not always done = “clinical trial”
 - “Controlled clinical trial” or “clinical trial” designs may have limited or no randomization
 - Eliminates bias (hopefully)
 - Random allocation vs. random selection (for surveys)
 - Random allocation: Subjects chosen for a research study are randomly placed (allocated) into one study group (intervention) or another (control, comparison, placebo)
 - Investigators usually define initial inclusion characteristics – define why certain subjects are included or excluded from the study overall
 - Random selection: People are randomly chosen (selected) from a group (population) to be in a research study
 - May not define specific inclusion criteria
 - Hidden bias introduced through imperfect randomization, failure to randomize all eligible patients, failure to blind assessors to patients’ randomization
- Groups receiving intervention(s) or control (comparison or placebo therapy) are identical (on average) with the exception of the intervention received. Differences in outcomes are attributable to the intervention only.
- **Strengths**
 - Strongest study design because there is so much control over the study
 - Investigators control the intervention
 - Allows rigorous evaluation of a single variable
 - Prospective: data is collected after the study is designed and in progress
 - Seeks to falsify (not confirm) its own hypothesis
 - Seeks to eradicate bias through comparison and blinding
 - Randomization decreases bias in group placement
 - Blinding of investigators to outcome measures decreases bias
 - Blinding more likely
 - Unbiased distribution of confounders: 2 or more factors that are “associated” (age and weight) and may affect (confuse, distort) the effect of the other(s) on the outcome (onset of diabetes)
 - Very structured therapy or intervention can be accurately described
 - Randomisation facilitates statistical analysis
 - Allows for “meta-analysis” (combining numerical results) at a later date.
- **Limitations**
 - Expensive and time consuming
 - Takes many research personnel to complete
 - May have limited applicability to a general population or a practice population due to tight inclusion and exclusion criteria
 - True randomization is difficult to achieve
 - Incomplete randomization
 - Volunteer bias
 - Bias in selection and randomization
 - Often impractical
 - Structured “design,” intervention, environment may be different than results a clinician would get in private (“real world”) practice
 - No variation from the intervention can be made by the research clinician
 - Ethically problematic at times
 - Other study designs may be more appropriate



Case-control studies

- Observational
- Focus on the etiology of a disease or health issue
- Patients with a particular health concern / characteristic / disease
- Matched with “controls:”
 - Identical patients without that issue
 - Identical patients with a different disease
 - General population
- Cases (with disease) vs. Controls (without disease) must be well-defined and adequately described
- Cases and controls should be taken from the same general population at risk of developing the disease, but with differing exposure to the potential risk factor. (minimizes selection bias)
 - Cases and controls should have the same opportunity to be exposed or receive the exposure.
 - Matching case and controls is one of the major challenges of this study design
 - Matching should go beyond demographics if other factors are known to be important for or affect the disease (e.g., general condition of health, ability to function, seek health care, etc.)
 - Can choose multiple control groups or multiple controls for each case); avoids selection bias (comparing 2 groups of patients who differ in more aspects than the one under study and one or more of those other aspects affect the outcome of the disease)
- Retrospective: look backwards in time
 - Data often is collected by searching through patient histories or through patient recall surveys
 - Compare past histories of possible risk factors between cases and controls
 - Best if the study involves new (incident) cases (less problems with history, recall)
- Used to study rare conditions (strong study design)
- Used to study the relative risk of disease related to a particular characteristic (genetic factor, exposure)
- Can be used to look at multiple factors or exposures for disease
- Validity depends on the ability to compare the case and controls
- **Strengths:**
 - Quick and not as complicated, expensive as RCT; short time required to conduct
 - Investigators can identify cases unconstrained by the natural frequency of the disease and are able to make comparisons
 - Only feasible method for very rare disorders or those with long lag between exposure and outcome (strong study design for rare diseases or diseases with a long latency)
 - Fewer subjects needed than cross-sectional or cohort studies
 - No additional risks to subjects (experimental interventions)
 - Existing records can be used (hospital records; health registries)
 - Can be used to study multiple factors affecting one outcome (disease)
- **Limitations:**
 - Case selection must be very well-defined (when is a “case” a “case”)
 - Matching should go beyond demographics if other factors are known to be important for or affect the disease (e.g., general condition of health, ability to function, seek health care, etc.)
 - Selection of control groups is difficult: people at risk of getting disease, but do not have the disease
 - Confounders: 2 or more factors that are “associated” (age and weight) and may affect (confuse, distort) the effect of the other(s) on the outcome (onset of diabetes)
 - Selection bias: investigators “create” the comparison groups rather than “letting nature take its course” in determining who in the population becomes a “case” and who remains a control.
 - Controls are not a “naturally occurring” group
 - Patients may differ in additional factors or aspects not under study that may affect the outcome of the disease
 - Measurement bias: exposure is measured after the onset of the disease or outcome under study; presence of outcome directly affects the exposure, affects subject’s recall of exposure or affects measurement or recording of the exposure.
 - Recall bias: Reliance on recall or records to determine exposure status (retrospective study)
 - Can only be used to study one outcome (disease)



Cohort Studies

- Observational
 - No randomization
 - No control over intervention or risk factor exposure (researcher observes subjects but does not control exposure)
- Provide a direct estimate of absolute risk: the probability of developing disease during a given time period
- Patients with similar characteristics at a common point in the course of the disease or health issue and followed over time using pre-defined measures of outcomes (pain, function, activities/quality of life, satisfaction with care, etc.)
- Prospective: follow groups forward in time from exposure to defined outcome of interest (disease)
- Measurement of the same outcome / issue
 - Patients suffering from low back pain
 - Death from heart attack
 - Subjects can be matched
- Two groups of patients differ in one characteristic
 - For example, smokers or non-smokers
 - Non-random allocation into one group or another (exposed / not exposed)
- Comparison group
- Eligibility criteria and outcome assessments can be standardized
- The best way to identify this study design:
 - Incidence rates defined
 - Natural history of disease is discussed

- **Strengths:**
 - Ethically safe;
 - Subjects can be matched; comparison group
 - Can establish timing and directionality of events
 - Eligibility criteria and outcome assessments can be standardized;
 - Administratively easier, less expensive, less complicated than RCT.

- **Limitations:**
 - Controls may be difficult to identify
 - Exposure may be linked to a hidden confounder;
 - Blinding is difficult
 - No randomization
 - Large sample sizes or long follow-up is necessary for rare disease
 - The expense and logistics associated with the attempt to compare the natural frequency of a potential disease associated with a particular exposure may not be feasible. A case-control study (fewer subjects needed) may be more appropriate.

- **Other names for Cohort Studies**
 - Incidence study
 - Longitudinal study
 - Forward-looking study
 - Follow-up study
 - Concurrent study
 - Prospective study



Case series

- Description of one group of patients (generally 10 or more) with similar diagnoses or therapy followed over time
- Descriptive study; does not test the hypothesis of treatment efficacy
 - Should not be used for comparison of treatments
- Should have:
 - Clearly defined question
 - Well-defined, detailed case definition
 - Very well-defined population involved in the study
 - Well-defined, well-described intervention, easily followed, replicated
 - Use of standardized descriptors, criteria and data
 - Use of validated outcome measures
 - Clear presentation of data and results
 - Appropriate statistical analyses
 - Larger number of cases (than a case study) allows statistical analysis (p values, means, standard deviations)
 - Well-described results focused on outcome measurement
 - Discussion and conclusions supported by data presented
 - Funding sources, affiliations acknowledged
 - IRB, human subjects review
- Multiple uses:
 - Case definition & detailed descriptors
 - Useful as a “benchmarking” descriptive study
 - Initial reports of new diagnosis or innovative treatment
 - Description of the natural history or natural progression of a condition or disease, recovery, complication rates
 - Trend analyses, descriptors, registry data of outcomes
 - Healthcare planning including economic analysis
 - Hypothesis, analysis of causation
 - Can be a hypothesis generating study, basis of follow-up studies
 - Multi-institutional registry
- All subjects receive same treatment
 - Treatment or intervention should be well described
- No comparison group
 - If inclusion and exclusion data were used, explicit definitions and descriptions should be provided
- Larger number of cases (than a case study) allows statistical analysis (p values, means, standard deviations)
- Allows determination of role of chance (as opposed to single case study)
- Often retrospective (look back in time) restricting value as prognosis study or determining cause and effect relationships
 - Prospective (looking forward) studies are often designed as prospective cohort studies, including a control group (a benefit, strength).
- **Strengths:**
 - Clearly defined question
 - Clearly defined study population
 - Well described study intervention
 - Outcome measures should be well-defined and validated
 - Well-described results supported by data and well-defined observations
 - Use of statistical analysis to assess the role of chance



Case series (con't)

➤ **Limitations:**

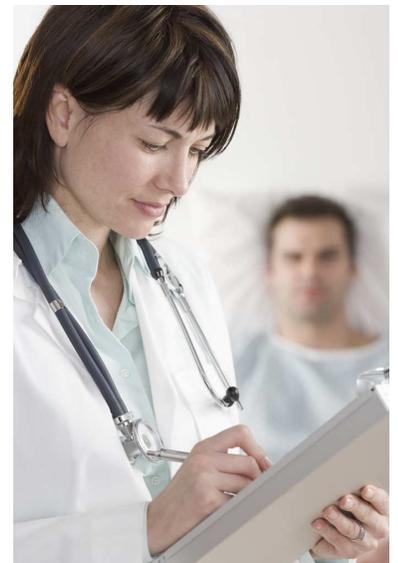
- No comparison group
- Blinding is unlikely
- Cannot be used to draw inferences regarding efficacy
- Not strong enough (typically) to test a hypothesis
- Published “benchmarking” studies usually are those with the best outcomes
- Study population may not be representative; generalization may be difficult
 - Study population may be too narrow to generalize to a different age, sex, culture, etc.
 - “Mixed” population may require larger sample sizes to realize trends in outcomes

Case report / case study

- Detailed description a single case
 - Describe rare events or early trends
 - Elucidate mechanisms of a disease or health issue and treatment
 - Describe unusual manifestations of a disease or health issue; describe an unusual response to an exposure or intervention
 - Highly detailed and methodologically sophisticated clinical and laboratory studies of a patient (small group of patients = case series)
 - Rich source of ideas, hypotheses about disease, conditions, risk, prognosis and treatment.
 - Not typically useful or strong enough to test a hypothesis
 - Initiate issues and trigger more decisive studies
 - Should have a very detailed, well-defined description of the patient
 - Do not include a statistical analysis; therefore, a determination of “chance” cannot be made
-
- **Strengths:**
 - Use as a “signal” to look for (or devise) further studies and evidence of the described phenomenon
 - All subjects receive treatment (at least some of the time)
 - Statistical tests assuming randomization can be used
 - Blinding can be maintained

➤ **Limitations:**

- Particularly susceptible to bias
- Not able to test most hypotheses
- Reports of successful therapy may be misleading since journals rarely print “negative” or unsuccessful case studies.
- Cannot be used to estimate the frequency of the described event (positive reaction of an intervention), role of bias or chance
 - Does not include a statistical analysis; therefore, a determination of “chance” cannot be made.



Cross-sectional surveys

- Representative sample of subjects or patients
- Interview, survey, study
- Data is collected at a single time point
- Data collection may depend on history or recall
- Establishes association, not causality
- Often used to develop further clinical research

- **Strengths:**
 - Less expensive and administratively simple
 - Ethically safe

- **Limitations:**
 - Establishes association at most, not causality
 - Recall bias susceptibility
 - Confounders (2 or more factors that are “associated” (age and weight) and may affect (confuse, distort) the effect of the other(s) on the outcome (onset of diabetes)
 - Neyman’s bias (incidence – prevalence bias or selective survival bias)
 - Group sizes may be unequal.



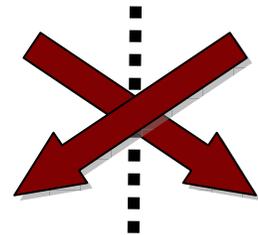
Cross-over Design

- Subjects are “moved” to the alternative group (intentional)
 - “Control” or placebo group receives treatment
 - Treatment group receives placebo or control treatment
- Intentional cross-over (by design) allows subjects to serve as their own control or placebo group.
 - Sample size is reduced (no need for an “equal set” of the control or treatment groups)
 - Error variance (statistical analysis) is reduced
- Unintentional cross-over between study groups (treatments and control) is often allowed for ethical reasons. However, studies should take into consideration the possibility of unintended cross-overs and allow for the possibility in the calculation of how many subjects are needed in a study.

- **Strengths:**
 - Intentional cross-over design is a very strong design, reducing variance among and between groups

 - All subjects receive treatment (at some point during the study)
 - Statistical tests assuming randomization can be used
 - Blinding can be maintained

- **Limitations:**
 - All subjects receive placebo or alternative treatment at some point
 - Washout period (treatment effect diminishes or ends) lengthy or unknown
 - Cannot be used for treatments with if the therapy (or control, comparison, placebo therapy) has permanent effects (subjects cannot “cross-over”)
 - Unintended cross-over (allowed for ethical reasons and patient preference) should be accounted for in calculations for the number of subjects needed for the study



Hierarchy of Study Designs



References:

Center for Evidence Based Medicine www.cebm.net EBM Tools: Study Design accessed December 12, 2007.

Fletcher RH, Fletcher SW, Wagner EH. Clinical Epidemiology, 3rd edition. Baltimore, MD: Williams & Wilkins 1996.

Greenhalgh T How to read a paper: the basics of evidence-based medicine (2nd Ed). London: BMJ Books, 2001.

Strauss SE, Richardson W S, Glasziou P, Haynes RB. Evidence-based Medicine: How to practice and teach EBM (3rd ed.). Edinburgh New York: Elsevier/Churchill Livingstone, 2005.

West S, King V, Carey TS, et al. Systems to rate the strength of scientific evidence. File Inventory, Evidence Report/Technology Assessment Number 47. AHRQ Publication No. 02-E0106, April 2002. Agency for Healthcare Research and Quality, Rockville, MD. Available at <http://www.ahrq.gov/clinic/streinv.htm>