Study designs: Intervention trials

Victor J. Schoenbach, PhD  www.unc.edu/~vjschoenb

Department of Epidemiology
Gillings School of Global Public Health
University of North Carolina at Chapel Hill
www.unc.edu/epid600/

This lecture will begin our consideration of the topic of epidemiologic study designs. Today we will present intervention trials – studies where the exposure of interest is an intervention by the researcher. We will also consider what is required to reach a causal conclusion, and how epidemiologic study designs try to approximate those requirements.
On inspiration

- A boy was watching his father, a pastor, write a sermon. “How do you know what to say?” he asked.
- “Why, God tells me.”
- “Oh, then why do you keep crossing things out?”
This lecture will cover:

• Causality - the counterfactual model
• Experimental and observational epidemiologic study designs
• Types of intervention trials
• Methodological issues
• Ethical issues
First, causality and the counterfactual model

How do we discover causation? In every day life we often “observe” causation by changing something and observing the result. Babies soon learn that if they let go of something it falls from their hands.

In his chapter on causal inference in *Modern Epidemiology*, Ken Rothman presents the example of a toddler playing with a lamp switch. She turns the switch, and the lamp lights. She turns the switch again, and the lamp goes out. She turns the switch again, and the lamp lights. Before long she knows that the switch turns the lamp on and off.

However, causation cannot be observed. Rather it must be inferred. The toddler is making an inference about the relation between the switch and the light. But inferences depend upon conceptual models, assumptions, and frames of reference. For example, the “cause” of the lamp lighting could be regarded as the mother’s having replaced the burned out light bulb earlier that day, or the paying of the electric bill, or the political and economic factors that induce electricity suppliers to construct generating facilities and make electricity available. A sufficient cause typically requires the action or presence of various component causes, some of which may be necessary causes.
The essence of the causal inference for the light going on is the comparison of two situations that we think differ by only one factor. In addition, the ability to change the factor, reverse the change, then change it again increases our assurance that no unseen factor was responsible. The modern formulation of the philosophy of causal inference in epidemiology, which derives from the philosophy of Hume, requires a comparison between (1) the outcome for people exposed to a factor (“the exposed”) and (2) the outcome that would have occurred if the same people had not been exposed, in order to infer causation. Alternatively, we can compare the outcome for people not exposed to the factor (“the unexposed”) with the outcome that would have occurred if the same people had been exposed. Since it is of course not possible to compare the same people with and without the exposure – at least not at the same time and with everything else the same – these comparisons are counterfactual. Only one condition can exist (e.g., the exposed condition); the comparison condition is counterfactual.
The closest that we can come to a counterfactual comparison in real life is a cross-over experiment, in which we apply the exposure, remove it, and perhaps apply it again. This strategy gives us nearly all of the inferential power of the counterfactual ideal, but it is available only for exposures that do not have lasting effects.

Thus most of the time we must compare outcomes in the exposed population to those for some other population, which we will call the “unexposed”. This unexposed population then serves as a substitute population. The idea is that outcomes for the substitute population should be the same as the outcomes in the counterfactual comparison. Thus, the substitute population represents the “exposed group without the exposure”. The validity of the inference we make depends on whether or not we can find a valid substitute population, a population whose outcome will be the same as the outcome for the exposed group if it had not had the exposure. (If we wish the index population to be the unexposed group, then the substitute population represents the unexposed population with the exposure.)
Most analytic epidemiologic studies of associations between risk factors and outcomes involve comparisons between “exposed” and “unexposed” groups. Most epidemiologic studies are observational in that the epidemiologist does not get to determine who becomes exposed. Sometimes the epidemiologist can observe the onset of exposure and then the development of the outcome. Most often, though, the exposed and unexposed groups exist before the epidemiologist arrives on the scene. In intervention studies, by contrast, the researcher decides who will be exposed. Intervention studies represent the experimental aspect of epidemiology. Even epidemiologic experiments are largely observational, since it is not possible to subject human research participants to the kinds of control and manipulations that can be done with laboratory animals and tissue cultures. However, the ability to assign exposure status is a crucial element, since it weakens other causal processes that could confound the comparison of exposed to unexposed. For this reason intervention studies have a special significance.
Analytic epidemiologic studies often follow one of four general study designs. Intervention trials are prospective studies in which an exposure or exposures are assigned to people and subsequent outcomes observed. Cohort studies are similarly prospective – they begin before the outcome has taken place and follow-up the study population to see who develops it. Cohort studies may or may not observe the onset of the exposure.

Many health outcomes occur in only a small proportion of the population, however, so it can be logistically challenging to observe a reasonable number of cases by following-up individuals. The case-control study offers an efficient way to study these rare outcomes, since it specifically enrolls people who experience the outcome and compares them to a sample of people representing the study base – the population from which cases arose.

The fourth major design is called cross-sectional. Cross-sectional studies measure the prevalence of conditions and characteristics at a point or during a period of time. The major periodic surveys conducted by the U.S. National Center for Health Statistics fall into this category.

Cross-sectional, cohort, and case-control studies are completely observational. Exposure status for individuals is not strongly influenced by the researcher and has often been determined before the researcher arrives on the scene. Moreover, any epidemiologic study may begin after a disease process has been already underway.
These general designs encompass many different types of studies. In the intervention arena, two major categories are therapeutic trials, involving treatment of people with a disease or other adverse condition, and preventive trials, seeking to prevent a disease from developing. An example of a major therapeutic trial is the Beta-blocker Heart Attack Trial (B-HAT). This study enrolled men who had just experienced a myocardial infarction (MI) – a “heart attack” – and randomly assigned them to receive a drug named propranolol and a placebo (an inactive preparation used to conceal which participants were actually receiving the medication). A “beta-blocker” is a drug that inactivates beta-adrenergic neurons*. The B-HAT trial was a success, in that participants receiving the active treatment were demonstrated to have better survival. The benefit became clear even before the trial had reached its planned duration, so it was stopped early. (Ref: The β-Blocker Heart Attack trial. JAMA 11/6/1981;246(18):2073)

* The mechanism of action of a neuron (a nerve cell) involves the binding of a chemical substance (neurotransmitter) to a specific type of molecule (“receptor”) in the cell’s outer membrane. One class of neurons respond to a neurotransmitter called norepinephrine (also once known as noradrenalin). There are two major types of receptors for norepinephrine – called alpha and beta. So beta-adrenergic neurons are nerve cells that respond to norepinephrine when that chemical binds to the beta variety receptor. Stimulation of beta-adrenergic neurons causes heart muscle to contract more vigorously and blood vessels to constrict, thereby increasing resistance to blood flow and blood pressure.
An example of a prevention trial is the trial of zidovudine (better known by its abbreviation, AZT) in the prevention of perinatal transmission of HIV virus from a pregnant mother to her infant. This trial is often referred to as "ACTG 076", because it was protocol #076 evaluated by the AIDS Clinical Trials Group (the term "protocol" is sometimes used to refer to a trial, especially one of a group of trials being conducted through a single organization or consortium). In ACTG 076, pregnant women who were infected with HIV were randomized to receive zidovudine orally beginning six weeks before their expected delivery date and intravenous zidovudine during labor and delivery. Their newborn infants were then given oral zidovudine until age six weeks. The trial was a great success, reducing the probability of maternal transmission of HIV by about two-thirds. As a result, all pregnant women are supposed to be counseled about and tested for HIV, with women who test positive being urged to take anti-retroviral medication (of which zidovudine was the first to be widely used). (For the early impact in N.C., see Susan A. Fiscus et al. Perinatal HIV infection and the effect of zidovudine therapy on transmission in rural and urban North Carolina. JAMA 1996;275(19):1483-1488.)

content.nejm.org/cgi/content/full/331/18/1173?ijkey=c98ee6dfa7368647a7a9b22ceaa0c26dd2f99c386
In my own career I have helped to lead two randomized trials of interventions to help people quit cigarette smoking. The first was Free & Clear, which was conducted among enrollees of Group Health Cooperative of Puget Sound, in Seattle, Washington, with Tracy Orleans (now at Robert Wood Johnson Foundation) and Edward Wagner, then Director of the Center for Health Studies at Group Health.

The *Quit for Life* study was a randomized trial of a self-help smoking cessation intervention designed for policyholders of North Carolina Mutual Life Insurance Company, then the largest African American insurance company in the U.S. Policyholders in 16 intervention districts received a packet of quitting aids from their NC Mutual sales agent, and their quitting experience was compared to policyholders in 16 control districts.

*Quit for Life* website: www.epidemiolog.net/pub/qfl/
I also This randomized trial was conducted to test whether having partner notification services for HIV-infected persons would lead to higher rates of notification of sexual and needle-sharing partners of people testing positive for HIV. The trial was conducted in three health departments in central North Carolina: Durham, Mecklenburg, and Wake.
The distinction between therapeutic and preventive trials is not always a clear one. Among the best known intervention trials are three major cardiovascular disease (CVD) prevention trials conducted during the 1970’s by the U.S. National Heart, Lung, and Blood Institute (NHLBI). The three trials were the Hypertension Detection and Follow-up Program (HDFP), Multiple Risk Factor Intervention Trial (MRFIT), and the Lipids Research Clinics Coronary Primary Prevention Trial (CPPT). The HDFP randomized women and men with “mild” hypertension to a systematic treatment regimen with a goal blood pressure. MRFIT randomized men with CVD risk factors placing them at high risk of coronary heart disease (CHD) to an intensive medical and behavioral intervention designed to reduce their risk. For these two trials, the control condition was “usual care”, that is, reports were provided to the men’s regular doctor, who was free to manage the patient as s/he wished. The CPPT randomized men in the top 15% of the distribution of blood cholesterol to take a cholesterol-lowering drug or a placebo. The objectives of these trials were to demonstrate that reducing risk factors could reduce risk of CVD, especially CHD. However, one can also consider the treatment of hypertension, hypercholesterolemia, and, perhaps tobacco addiction, as treatment of disease.

There have been several community trials in Tanzania and Uganda to test whether treatment of bacterial sexually transmitted infections can reduce risk of HIV transmission. A summary of three of these trials, which had conflicting results, can be found in Blackboard.
Intervention studies can be subcategorized in many ways. Smaller studies are conducted from a single site. For example, the *Free & Clear* intervention trial delivered an intervention and collected data by mail and telephone from the headquarters of Group Health Cooperative in Seattle. In order to enroll an adequate number of participants, many intervention trials involve multiple sites, which may also have their own investigators. In such cases a steering committee and a coordinating center develop a detailed protocol that all sites are required to adhere to, manage the data centrally, and have review mechanisms regulating use of the data. The *Quit for Life* study was carried out by NC Mutual sales agents in various sales districts, but there was only a single investigator team.

Multisite trials present challenges for coordination and standardization. Both single-site and multisite intervention trials carried out in resource-poor settings present additional challenges. Infrastructure (clinics, laboratories, equipment) may need to be developed, recruitment of staff with required expertise and experience may be difficult, multiple ethical and regulatory approvals may be required, and relationships with communities and political leaders may require special emphasis. See Jennifer Deese’s account of “Challenges of International Intervention Trials”, in the Supplementary Materials folder in Blackboard || Intervention Studies.
Another classification of intervention trials is whether they apply the intervention to individuals or to communities. Individual-level interventions, such as medication, psychotherapy, or counseling are often referred to as “clinical” trials, though participants may be students, workers, or other persons who are not “patients”.

In community trials, communities are assigned, ideally randomly, to intervention or control conditions, and the intervention is applied at the community level. Group-randomized trials are often used for interventions where the intervention, such as an educational program or an environmental alteration, cannot readily be restricted to particular individuals.

Individual-level trials are typically more statistically efficient – they provide more statistical power per person enrolled – and can permit greater control over the intervention, though that may be more a matter of the type of intervention that can be tested in an individual-level trials (e.g., a drug). Group randomized trials have important advantages because they are closer to the scale of a public health intervention. However, the statistical implications of randomizing groups rather than individuals can make these trials much larger than would be the case if the intervention could be randomized at the individual level.
A crucial design feature of intervention trials is whether the intervention is assigned at random or not. Randomization is a key feature of study design.
Randomization

Why is randomized assignment of intervention so important?

Before proceeding, see what reasons you can come up with for the importance of randomization. Can you identify at least three?
Randomized assignment is so important because it provides the strongest evidence for causal inference. A randomized trial makes the claim that the group differences in outcomes reflect the intervention and not investigator manipulation more credible. Here is why.
Thinking back to the conceptual framework for causal inference presented in the introduction to this lecture, we can see that a randomized control group is the most likely to be a valid substitute population. The desired comparison for exposed persons is the counterfactual, the exposed persons without the exposure. Randomization provides a substitute population that comes at least on average the closest to this counterfactual comparison group.

So the first reason for the importance of randomization is that it provides the best assurance that the control group is a valid substitute population. Randomization avoids self-selection for the exposure as well as provider selection that might produce non-equivalent groups.
Another reason, which may seem abstract at this point in the course, is that randomization is the only means in epidemiology to control for the influence of factors that are unknown or unmeasurable. Many outcomes are influenced by multiple factors. When we study the influence of one factor we obviously need to take into account, or control for, the influence of other factors, in order to isolate the relation of interest. But except for randomized assignment of exposure, the various strategies available for doing this require that we be able to identify and measure these factors.

### Why is randomized assignment of intervention so important?

1. Best assurance that control group (unexposed) is a valid substitute population

2. Only way to control for unknown factors
Yet another reason for the power of randomization is that it facilitates masking the participant, the observer, and the analyst concerning the participant’s true exposure status. Accuracy of measurement and classification are central concerns in research. Of particular concern is the possibility that exposure status may influence the classification or measurement of the outcome, since such differential classification could readily distort the apparent relation between exposure and outcome. In most cases, preventing everyone concerned (participant, observer, analyst – triple masking) from knowing participants’ true exposure status ensures that the measurement of the outcome and the analysis of the data will be even-handed with respect to the exposure.
Yet one more reason, though this is often true of nonrandomized intervention studies as well, is that randomized assignment avoids ambiguity concerning the time order of the exposure and the outcome. Observational epidemiologic studies must often struggle with uncertainty in this respect, since disease may go unrecognized for a prolonged period and may even influence exposures before the disease becomes recognized. For example, some diseases may influence appetite and digestive comfort, leading to changes in diet and/or use of medications. When the disease becomes manifest, it could appear that the diet or medicine is associated with the disease (e.g., use of anti-ulcer medication and stomach cancer). With randomized treatment assignment, there is the assurance that the intervention could be given differentially to people with occult disease only by chance. Although even a chance association can mislead us, at least we can quantify the probability of such an occurrence.
A methodological issue that the biostatisticians among you will appreciate is that randomization provides the foundation for statistical tests and measures of precision. Randomization enables valid quantification of uncertainty. The theory that underlies all common statistical tests and variance estimation procedures assumes random processes with certain characteristics. In most epidemiologic studies of exposure-disease associations, the assumptions behind statistical tests and estimation procedures are not in fact satisfied, leaving in doubt the validity of standard statistical methods. Randomized assignment alone satisfies the assumptions behind statistical significance tests in comparative studies.
All research involves attempts to control unwanted influences so that relations observed can be attributed to the factors under study. In laboratory experiments the researcher can gain a great deal of control by specially designing and supervising the environment in which the experiment is carried out. The benefits described for randomization require that it be strictly applied and in general that there be strict adherence to the protocol of procedures for intervention and observation. However, it is difficult to exercise control, especially over people, as those of us who have parents, spouses, children, bosses, or subordinates are acutely aware. But because intervention trials come the closest to true experiments, the impetus to control what happens is probably the strongest in these studies. An epidemiologic study is an exercise in management, and this is particularly the case in intervention studies.
In a randomized trial it is not only the research participants and the data collectors whom the investigators strive to control. They also strive to control themselves. A randomized trial is a true experiment, often a very expensive one that is undertaken in order to “prove” a hypothesis and influence policy. As mentioned above, a randomized trial can produce much stronger statistical evidence than non-randomized studies. In order to preserve these strengths, biostatisticians have prominent roles in major randomized trials and attempt to enforce classical statistical hypothesis testing procedures developed by Neyman and Pearson. The biostatisticians will insist that the primary analytic procedures be specified in advance of the study, including the specific hypothesis, the primary outcome variable, the primary statistical test to be used to determine the difference in outcomes between intervention and control groups, and the rules for deciding whether and when the trial should be stopped before its planned length.

The primary statistical analyses will generally use an intention-to-treat approach. Intention-to-treat classifies participants as “exposed” if they were originally assigned to the exposure group, even if they did not receive the intervention (e.g., due to side effects or noncompliance). Similarly, participants assigned to the control group are classified as “unexposed” even if they received the intervention outside of the trial. Despite the problems that arise from such misclassification, maintaining the original group assignment is necessary in order to preserve the statistical benefits of randomization.
Consideration of early stopping is generally required for ethically reasons. The ethical basis for randomized trial is what is called “equipoise” – genuine uncertainty about whether the intervention produces a better outcome. Once the intervention has been clearly demonstrated to be superior, it is generally not ethical to withhold it from the control group. Also, if the intervention is found to be harming those who receive it, it must be stopped as soon as possible.

In order to avoid prematurely terminating a trial because of random fluctuations in outcomes, statistical procedures are used to require a higher threshold to stop a trial early than to find a significant difference at the planned termination date.
So, “listen to your statistician”

In the same vein, clinical trials must now be prospectively registered in a public database, such as clinicaltrials.gov

Avoids the problem of “negative” studies becoming buried in file drawers

So listen to your statistician.

Also, clinical trials must now be prospectively registered in a public database, such as clinicaltrials.gov
The Coronary Primary Prevention Trial (CPPT), originally called the Lipid Research Clinics (LRC) trial, illustrates these principles (and, in fact, the LRC Coordinating Center has been led by biostatisticians – first O. Dale Williams, then C. Edward Davis, and now Woody Chambless, all of the UNC-CH Department of Biostatistics). The CPPT tested the hypothesis that reducing serum cholesterol through the use of the drug cholestryramine would reduce CHD incidence. CHD was measured with clearly specified diagnostic criteria. The primary analysis compared CHD incidence rates between cholestyramine and placebo groups regardless of compliance. Interim analyses were carried out with a procedure that required a much stronger difference in order to stop the trial early but reduced statistical power for the final comparison only slightly.

In fact, the strategy worked. The trial demonstrated the hypothesized effect with the specified outcome and statistical test and was regarded by many as providing definitive proof that lowering cholesterol prevents CHD. The overall mortality reduction in the treatment group was not statistically significant, but the trial was not designed to show such a difference. The National Cholesterol Education Program was launched following the announcement of the CPPT results (which must have pleased the researchers involved more than did the launch of the National High Blood Pressure Control Program before the conclusion of the HDFP). Although the CPPT studied only middle-aged men with the highest cholesterol levels, the results were then extrapolated to women, younger persons, and to those with moderately high cholesterol levels.
A strong statistical demonstration provides powerful evidence for the hypothesized relation. However, other evidence is also necessary to make a convincing case that the intervention caused the difference in outcomes. For example, no matter how strong the statistical results, how can one infer that an exposure that was equally present in intervention and control groups was responsible for a difference in their outcomes? Thus it is necessary to establish that the intervention group, or at least most of the intervention participants, did indeed receive the intervention, and that most of the control participants did not.
So it is essential to verify that the intervention and control groups did indeed differ in regard to their exposure to the intervention. The twin concerns are 1) compliance – did intervention participants receive the intervention, and 2) unplanned crossovers – did control participants receive the intervention, for example, from outside the trial.
The term “crossover” refers to participants in one experimental condition having the exposure status intended for another experimental condition. This crossing over happens by design in a crossover study, where the investigator arranges for all participants in one condition to then receive the exposure for the other condition, and vice-versa. As noted earlier, such planned crossing over, when the exposure and outcome make it a possibility, can provide the closest approximation to a counterfactual comparison.

Unplanned crossovers are an entirely different matter. Unplanned crossovers occur when participants do not have the exposure status intended for the experimental condition to which they were assigned. If the exposure truly affects the outcome, then unplanned crossovers will most likely weaken the effect that will be observed. If the number of unplanned crossovers is sufficiently large, they can completely obscure the true effect. At the same time ethical and practical considerations limit what investigators can do to ensure that participants stay in their assigned treatments.
To demonstrate the difference in exposure, we need to measure it. Although measures of exposure must often rely heavily on what the participant says, pill counts, analysis of biological specimens, and behavioral observation may be possible. Data from the same or similar procedures with the control participants (if, for example, they are taking placebo pills) can be useful for analysis and interpretation. Such data have shown that compliant controls (participants in the control group who took their placebos) often have better outcomes than noncompliant controls, which suggests the importance of the intent-to-treat analysis.
As noted, the primary analysis will usually be based on intention-to-treat, in order to preserve the inferential advantages of randomized assignment. A treatment-received analysis should also be conducted. In a treatment-received analysis, participants are categorized according to their actual exposure to the intervention. The results of the two analytic approaches – intention-to-treat and treatment-received – should be similar. In principle the treat-received analysis should produce a stronger association, since there is less misclassification of exposure status. However, because participants in each exposure group are to some extent “self-selected”, other biases could change the association in either direction.
Besides demonstrating that the intervention and control groups did actually differ in their exposure status and that exposure status was in fact associated with outcome, we also want to show that the relation between exposure and outcome came about by the hypothesized mechanism. Thus we will want to measure intermediate variables that are believed to mediate the effects of the intervention on the outcome. Data on these variables are used to demonstrate that the effect occurred by the hypothesized mechanism – or, if the effect was not observed, to help in understanding why not. In the CPPT, the drug (cholestyramine) was supposed to prevent CHD by lowering serum cholesterol levels. Thus it was important to show that in fact cholesterol levels in the intervention group declined, and that they declined more than in the control group (some decline was supposed to occur in both groups, because all participants were instructed in a cholesterol-lowering diet). In fact, cholestyramine compliers and the intervention group as a whole had a greater reduction in cholesterol levels. Analysis of the relation between cholesterol levels across participants and CHD incidence observed a 2% reduction in CHD incidence for every 1% reduction in cholesterol level. This relation confirmed the expectation based on earlier, observational studies.
The inferential strengths of intervention trials do not make them immune to criticism (naturally). A dilemma in intervention trials is that the evidence they provide is often strongest when participants are more highly selected, so that their risk is sufficiently high to ensure a sufficient number of outcome events, variability in factors that might confuse the comparison is reduced, and unplanned crossovers are minimized. For this reason, participants often comprise a small proportion of the population for whom the results may be relevant. The question is then raised, to whom can these results be generalized? The issue is one of “epidemiologic generalization”, rather than statistical generalization. The participants do not represent a random sample from the broader population, since they were selected to have specific characteristics (or they selected themselves).

Participants in the CPPT had cholesterol levels in the top 15% of the population distribution. Participants were all middle-aged men. Is it appropriate to generalize the findings to other groups? If not, how feasible is it to conduct similar studies with participants from these other groups. For that matter, even if a trial is not highly selected, so that all groups are represented among its participants, do the findings apply to all these population groups? Statistically, the results apply only to the source population as a whole unless the relation exists in each of the corresponding subgroups in the study. Generalizability is often a challenging question in population research.
Occasionally results of a randomized trial are in conflict with results from observational studies that preceded it. This situation arose with the results of the Women’s Health Initiative, a major study conducted to establish whether hormone replacement therapy (HRT) – estrogen (plus progesterone in women with an intact uterus) would reduce cardiovascular disease risk and several other outcomes. Many observational studies – case-control studies, cohort studies – had supported the hypothesis that hormone replacement therapy reduced CVD risk in postmenopausal women, and many women were receiving prescriptions. But some scientists argued that a randomized trial was essential to demonstrate the relation, and were eventually able to persuade other scientists who felt it would be unethical to conduct a trial in which some women would receive a placebo rather than this beneficial treatment.

So the Women’s Health Initiative (WHI), a huge study, was designed to test HRT in relation to several outcomes including CVD and breast cancer. As it turned out, the trial was stopped early because the data indicated that women taking hormone replacement therapy had higher rates of CVD and breast cancer.

Several theories were advanced about why the WHI did not find the benefit that had been seen in the observational studies, such as that the benefit required that women receive HRT at the time of menopause, not starting only years later. But none of the alternative explanations have held up. So the WHI results changed the practice of medicine.
Finally, intervention studies often raise difficult ethical issues. The ethical basis for an intervention trial is that the scientific community is in equipoise – it is not possible to say whether the intervention is in fact better than the alternative. There is a delicate balancing between having enough evidence in favor of the intervention and having too much. Unless there are strong indications that the intervention will be beneficial, then why expose people to it and why devote the resources to testing it. But if the evidence is very strong, it may not be ethical to withhold the treatment from a comparison group. So there must be genuine uncertainty about which treatment is better.

A particularly difficult issue concerns intervention trials to test something which is known not to be the best available treatment but which may be more available, for example, because it is less expensive. Several trials of alternatives to the 076 zidovudine regimen to reduce maternal transmission of HIV were carried out in Thailand and countries in Africa. Even though these trials had been accepted by the countries where they were being conducted, and in some cases even requested by the country, the trials provoked intense controversy.

For any research project, and especially an intervention trial, the question increasingly arises about who will benefit from the knowledge gained. Often the intervention is provided to the control participants, if it will still be beneficial. But what if the fruits will not be accessible to the entire community?
Adaptive designs

- Early results from a trial can be used to adjust the allocation odds so that more patients receive the apparently better treatment (e.g., “randomized play-the-winner”)
- Example: pregnant women randomized to AZT or placebo; adaptive design would have resulted in fewer HIV+ infants

Connor EM, Sperling RS, Gelber R, et al. (Reduction of maternal-infant transmission of human immunodeficiency virus type 1 with zidovudine treatment. New Engl J Med 1994;331:1173–1180) used standard randomization to administer AZT to 239 pregnant women and placebo to 238. Results: 20 HIV+ newborns in AZT, 60 in placebo. A randomized play-the-winner trial would have required more subjects for the same power, but even so would have fewer HIV+ newborns. (William F. Rosenberger, Randomized Play-the-Winner Clinical Trials: Review and Recommendations. Controlled Clinical Trials 1999;20:328–342)
And, as a little gift for your patience in listening to the entire lecture, here is a little story about a water pistol.

When a 3-year-old boy opened his birthday gift from his grandmother, he discovered a water pistol. He squealed with delight and headed for the nearest sink.

The boy's father was not so pleased. He turned to his mother and said, “I'm surprised at you. Don't you remember how I used to drive you crazy with water guns?”

Mom smiled and then replied … “I remember.”

The Water Pistol

When a 3-year-old boy opened his birthday gift from his grandmother, he discovered a water pistol. He squealed with delight and headed for the nearest sink.

The boy's father was not so pleased. He turned to his mother and said, “I'm surprised at you. Don't you remember how I used to drive you crazy with water guns?”

Mom smiled and then replied … "I remember."