Yarmouk University

Community Medicine

Lec. 15 - Experimental Design

Written By : Rahma Marie & Abdallah AlKashi

If you come by any mistake , please kindly report it to shaghafbatch@gmail.com

Text in Green Color is From Record

Assessing Preventive and Therapeutic Measures: Randomized Trials

DR. REEMA KARASNEH

COMMUNITY MEDICINE (MED 410)

Learning Objectives

- To describe the important elements of randomized trials.
- To define the purpose of randomization and of masking.
- To introduce design issues related to randomized trials, including stratified randomization, planned and unplanned crossovers, and factorial design.
- To illustrate the problems posed by noncompliance in randomized trials.

Introduction

- The objective, both in clinical practice and in public health, is to modify the natural history of a disease so as to prevent or delay death or disability and to improve the health of the patient or the population.
- The challenge is to select the best available preventive or therapeutic measures to achieve this goal.
- To do so, we need to carry out studies that determine the value of these measures.
- The randomized trial is considered the ideal design for evaluating both the efficacy and the side effects of new forms of intervention.

Sheet# 1

- The evidence needs to be extraordinarily strong to be able to call a factor "a preventive agent" for a disease.
- When we talk about diseases, most of the preventive measures or treatments are going to be medications. So, we can check their efficacy and side effects. We can know if the drugs are effective to the prognosis of the disease based on the goal we have in mind. How do you evaluate the efficacy and side effects? You get the values from randomized trials.
- The investigator introduces an intervention to the participant or patient. Randomized trials can also be used for screening. For example, does screen for prostate cancer reduce the mortality in this population? Have two groups: a group to perform prostate screening on and a control group. Is there a difference in mortality between them?
- Anything with an intervention can be evaluated with randomized trials.

Randomized trials

- Trials are essentially experiments which are under the control of the investigator.

- Randomized trials can be used for many purposes.
  - Evaluating new drugs and other treatments of disease, including tests of new health and medical care technology.
  - To assess new programs for screening and early detection
  - to compare different approaches to prevention, or new ways of organizing and delivering health services

RANDOMIZED CLINICAL TRIALS

- Compared to all other study designs, randomized clinical trials offer the most compelling evidence of "cause and effect."
- Employ a formal mechanism of chance to assign participants to receive an intervention of interest versus a control.
- Then, subjects followed over time to measure one or more outcomes.

Sheet# 2

- Randomized trials offer a chance to compare between people who receive an intervention and those who do not.
- Subjects can be followed over time to see whether an outcome has happened or not. Outcomes can be:
- Hypertension medication (intervention) leads to an outcome of controlled blood pressure.
- Outcome can also be survival. If survival decreases, you stop the trial.

- Case control and cohort are observational only. You do not offer intervention. The only similarity between them is the comparison between groups. Randomized trials offer intervention. If you make a new drug you have to offer it to participants to compare treatment vs no treatment.

The basic design of a randomized trial

- Begin with a defined population in which participants are randomized to receive either a new treatment or the current treatment.
- Then follow the subjects in each group to see how many are improved in the new treatment group compared with how many are improved in the current treatment group (often referred to as "usual care" or "standard of care").
- If the new treatment is associated with a better outcome, we would expect to find better outcomes in more of the new treatment group than the current treatment group.

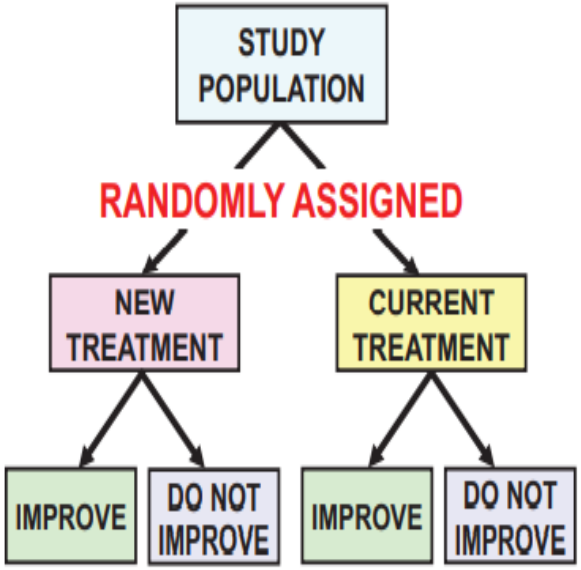


Fig. 10.1 Design of a randomized trial.

Sheet# 3

- The basic design of randomized trial: you have a study population; you randomly assign some of the participants to take the new treatment while the others continue using the old treatment. Then you compare the improvement between the groups to decide if the new treatment is effective.

- We may choose to compare two groups receiving different therapies, or we may compare more than two groups.

- Although at times a new treatment may be compared with no treatment, often a decision is made not to use an untreated group.
  - For example, if we wanted to evaluate a newly developed therapy for acquired immunodeficiency syndrome (AIDS), would we be willing to have a group of AIDS patients in our study who were untreated? The answer is clearly no;
  - we would compare the newly developed therapy with a currently recommended regimen, which would clearly be better than no therapy at all.

Sheet# 4

- You can also compare between more than two groups. For example, group A will continue using the old treatment, group B will undergo no treatment at all, and group C will undergo the new treatment.
- Usually, you compare between people using an old treatment vs new treatment. Using people who are not undergoing any kind of treatment (group B in the earlier example) is not preferred. Why? Because you are harming them while waiting to see the efficacy of the new treatment.
- Using a control group using a placebo is unethical unless you are not harming them.

Issues that must be considered in the design of randomized trials.

- Specification of the study "arms," or treatments.
  - These must be clearly stated
- Criteria for the measurement
- The duration of the treatments
- How long the study will last

Sheet# 5

- Specification of the study arms: arms are groups. Control is an arm, group using the new treatment is another arm. So, you must specify who is using the new treatment vs those who are not. Each arm must have specific qualities to decide what participants will go to which arm. To improve the comparison, the arms must be similar in qualities expect for taking the treatment we are studying.
- Criteria for measurement: what are you measuring? Improvement? How? By comparing survivability or mortality.

Selection of Subjects

- The criteria for determining who will or will not be included in the study must be spelled out with great precision and in writing before the study is begun.
- Any study procedure must in principle be replicable by others.
- Clearly, this is easier said than done because in randomized trials we are often dealing with relatively large populations.

Sheet# 6

- Selection of subjects: when writing the method of the study, the reader must be able to replicate the study and achieve the same results. So, selection of subjects based on specifications must be clear so the study can be replicated in the future. Because of the large populations in randomized studies, precision can be difficult because different people have different qualities. A range of variability is allowed.
- There are phases in randomized trials. With each phase, the population sample increases in size. You start with the smallest number possible and increase with each phase as you measure how safe the treatment is. If the intervention is something like screening or education, you can include large population sizes from phase 1 because there are no harmful side effects or risks.

Allocating Subjects to Treatment Groups Without Randomization

- STUDIES WITHOUT COMPARISON
- STUDIES WITH COMPARISON
  - Historical Controls
  - Simultaneous Nonrandomized Controls

STUDIES WITHOUT COMPARISON

- In this type of study, no comparison is made with an untreated group or with a group that is receiving some other treatment.
- The issue of comparison is important because we want to be able to derive a causal inference regarding the relationship of a treatment and subsequent outcome.
- There is a problem of inferring a causal relationship from a sequence of events without any comparison
  - E.g. If we administer a drug and the patient improves, can we attribute the improvement to the administration of that drug?

Sheet# 7

- Studies without comparison: this design is used but you cannot decide if there is a cause-and-effect relationship because there is no comparison. It can be used to see the side effects of a treatment instead of efficacy. Due to its flaws, we use comparison design (basic design).

STUDIES WITH COMPARISON: Historical Controls

- We could use a comparison group from the past, called historical controls.
  - E.g. We have a therapy today that we believe will be quite effective, and we would like to test it in a group of patients; we know that we need a comparison group. So, for comparison, we will go back to the records of patients with the same disease who were treated before the new therapy became available.

- This type of design seems inherently simple and attractive. But???

Problems in using historical controls

1. **Difference in the quality of the data collection.**

- We may set up a very meticulous system for data collection from the patients currently being treated. But we cannot do that for the patients who were treated in the past, for whom we must abstract data from medical records which are likely useful for managing individual care but are fraught with error and omissions when used for research purposes.
- Consequently, if at the end of the study we find a difference in outcome between patients treated in the early period (historical controls) and patients treated in the later (current) period, we will not know whether there was a true difference in outcome or whether the observed difference was due only to a difference in the quality of the data collection.
- The data obtained from the study groups must be comparable in kind and quality

Problems in using historical controls

2. **"Secular changes."**

- If we observe a difference in outcome between the early group and the later group, we will not be sure that the difference is due to the therapy because many things other than the therapy change over calendar time (e.g., ancillary supportive therapy, living conditions, nutrition, and lifestyles).

- Hence, if we observe a difference and if we have ruled out differences in data quality as the reason for the observed difference, we will not know whether the difference is a result of the drug we are studying or of other changes that take place in many other factors that may be associated with the outcome over calendar time.

Sheet# 8

- Historical controls: you have a treatment, and you offer it to a group, and you followed up over time. Where do you get controls? You can go back to hospital records to find patients with the same disease who did not take the treatment. Observe the progression of disease and prognosis. So, there is comparison in historical controls.
- It seems simple but the problem is:
- 1) some factors could have existed between the two groups due to the difference in time periods
- 2) some habitual differences could have existed between the two groups due to the difference in time periods
- 3) there will be a difference between the data collection quality. The current group with the treatment had data collected for the purpose for research. The historical control group picked from past medical records had data collected for medical purposes.

So,…When historical controls are useful?

- When a disease is uniformly fatal and a new drug becomes available, a decline in case-fatality that parallels use of the drug would strongly support the conclusion that the new drug is having an effect.
  - **Examples include the discovery of :**
    - insulin to treat diabetes
    - penicillin to treat serious infections
    - tyrosine kinase inhibitors (TKIs) such as imatinib (Gleevec) to treat chronic myelocity leukemia.

We use historical controls during the covid pandemic for example.

- Nevertheless, the possibility that the decline could have resulted from other changes in the environment would still have to be ruled out.
- Other changes in environment include lifestyle changes during the different time periods.

STUDIES WITH COMPARISON: Simultaneous Nonrandomized Controls

- To use simultaneous controls that are not selected in a randomized manner.
- **There are a number of possible approaches for selecting controls in such a nonrandomized fashion:**
  1. **Day-of-the-month method:** To assign patients by the day of the month on which the patient is admitted to the hospital:
     - For example, if admission is on an odd-numbered day of the month the patient is in group A, and if admission is on an even-numbered day of the month the patient is in group B.
  - The problem here is that the assignment system was predictable:
    - it was possible for the physicians to know what the assignment of the next patient would be.
- The goal of randomization is to eliminate the possibility that the investigator will know what the assignment of the next patient will be, because such knowledge introduces the possibility of bias on the part of the investigator regarding the treatment group to which each participant will be assigned.

Sheet# 9

- Simultaneous nonrandomized controls:
- You have a group that took the treatment and a control group. You either pick them in randomized fashion or nonrandomized fashion.
- Nonrandomized: source of selection bias. And source of observer bias because the investigator can predict the participant belongs to which group and can pick the outcomes he wants.

Allocating Subjects Using Randomization

- Randomization is the best approach in the design of a trial.
- Randomization means, in effect, tossing a coin to decide the assignment of a patient to a study group.
- The critical element of randomization is the unpredictability of the next assignment.
- Randomization removes selection and observer bias.

How is randomization accomplished?

1. **Computer programs**
2. **Manual randomization**
   - Table of Random Numbers

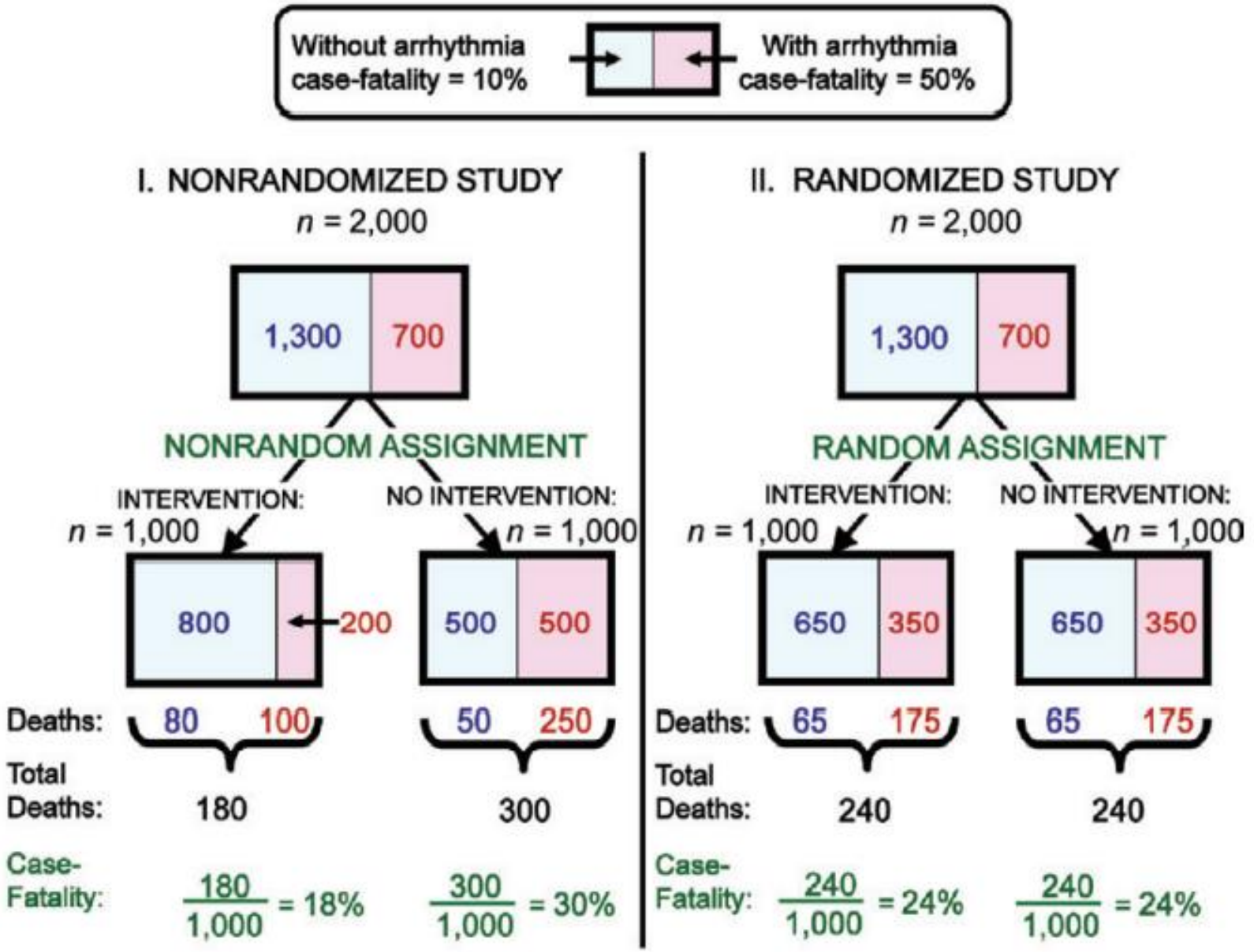


Fig. 10.3 Nonrandomized versus randomized studies. *I*, If the study is not randomized, the proportions of patients with arrhythmia in the two intervention groups may differ. In this example, individuals with arrhythmia are less likely to receive the intervention than individuals without arrhythmia. *II*, If the study is randomized, the proportions of patients with arrhythmia in the two intervention groups are more likely to be similar.

Example: Fig. 10.3

- Fig. 10.3 presents a hypothetical example of the effect of lack of comparability on a comparison of mortality rates of the groups being studied. Let us assume a study population of 2,000 subjects with myocardial infarctions, of whom half receive an intervention and the other half do not.
- Let us further assume that of the 2,000 patients, 700 have an arrhythmia and 1,300 do not. Case-fatality in patients with the arrhythmia is 50%, and in patients without the arrhythmia it is 10%.
- Let us look at the nonrandomized study on the left side of Fig. 10.3. Because there is no randomization, the intervention groups may not be comparable in the proportion of patients who have the arrhythmia.
- Perhaps 200 in the intervention group may have the arrhythmia (with a case-fatality of 50%) and 500 in the no-intervention group may have the arrhythmia (with its 50% case-fatality). The resulting case-fatality will be 18% in the intervention group and 30% in the no-intervention group. We might be tempted to conclude that the intervention is more effective than not intervening.

Example Fig. 10.3 - Cont.

- However, let us now look at the randomized study on the right side of Fig. 10.3. As seen here, the groups are comparable, as is likely to occur when we randomize, so that 350 of the 1,000 patients in the intervention group and 350 of the 1,000 patients in the no intervention group have the arrhythmia.

- When the case-fatality is calculated for this example, it is 24% in both groups. Thus the difference observed between intervention and no intervention when the groups were not comparable in terms of the arrhythmia was entirely due to the noncomparability and not to any effects of the intervention itself.
- (Please note that although Fig. 10.3 shows 1,000 participants in both the intervention and no-intervention group, randomization does not guarantee an equal number of participants in each group; however, with large numbers, on average the two groups will generally be comparable.)

What Is the Main Purpose of Randomization?

1. **To prevent any potential biases on the part of the investigators from influencing the assignment of participants to different treatment groups.**
   - When participants are randomly assigned to different treatment groups, all decisions on treatment assignment are removed from the control of the investigators. Thus, the use of randomization is crucial to protect the study from any biases that might be introduced consciously or subconsciously by the investigator into the assignment process.

2. **The randomization often increases the comparability of the different treatment groups; However, randomization does not guarantee comparability.**

3. **To whatever extent randomization contributes to comparability, this contribution applies both to variables we can measure and to variables that we cannot measure and may not even be aware of, even though they may be important in interpreting the findings of the trial**.

Sheet# 10

- Sometimes, you do not take into account specific factors. Randomization contributes to comparability and takes those factors into account as well.

Stratified Randomization:

- Sometimes we may be particularly concerned about **comparability** of the groups **in terms of one or a few important characteristics** that we strongly think may **influence prognosis** or **response to therapy** in the groups being studied, but as we have just said, randomization does not ensure comparability.

- An option that can be used is **stratified randomization**, an assignment method that can be very helpful in increasing the likelihood of comparability of the study groups.

Sheet# 11

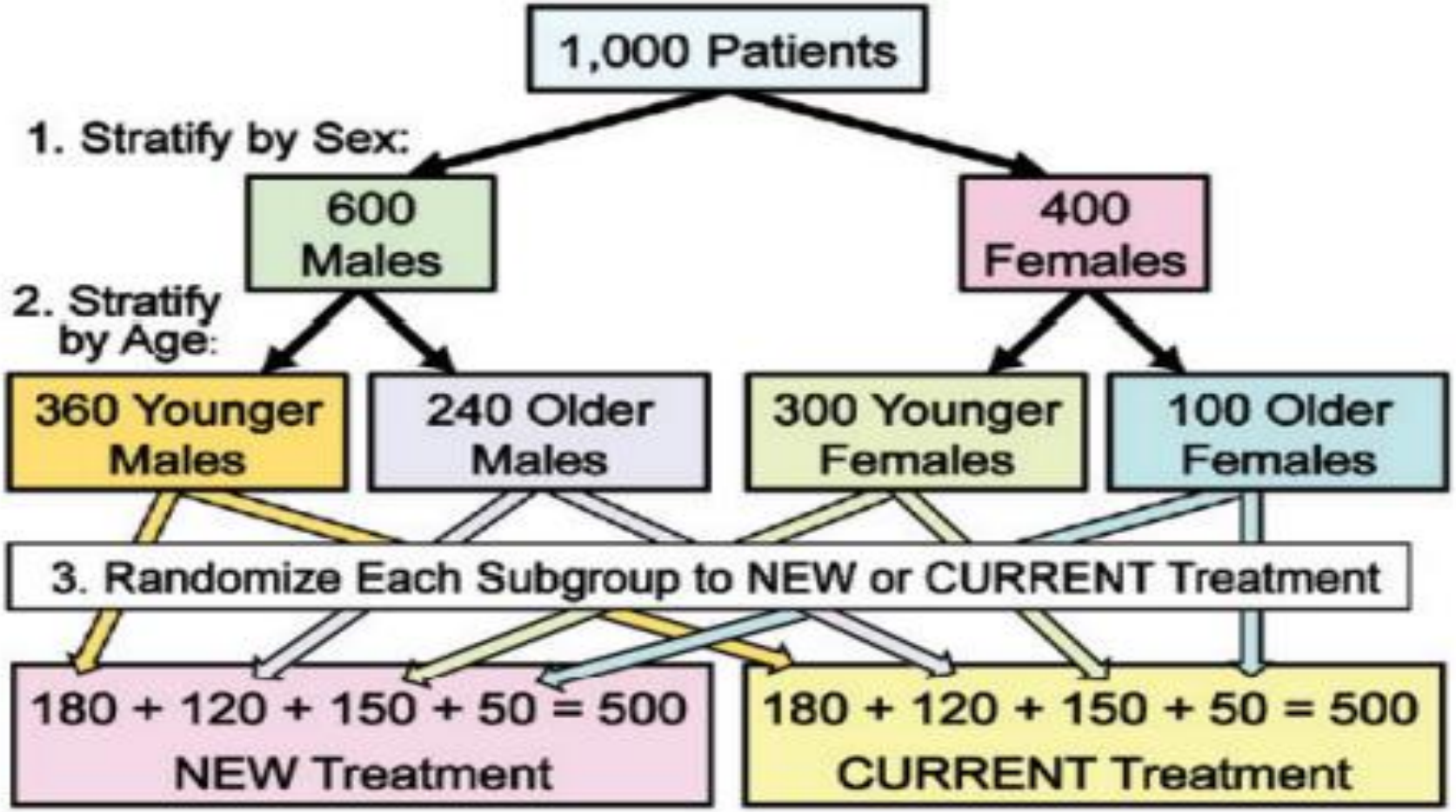
- Stratified randomization:
- Improves comparability by taking into account more than one suspected characteristic that might influence the prognosis. For example, gender and age wise. First, we arrange them based on age and gender, then we form the sample randomly. Now we have equal genders and ages in both groups. Now we are sure of the comparability من ناحية الجنس والعمر.
- Example: Refer to slide 37 for the example.

How stratified randomization is used to assign participants to different study groups?

- For example, let us say that we are particularly concerned about **age as a prognostic variable**: prognosis is much worse in older patients than among the younger.
- Therefore, we are concerned that the two treatment groups be **comparable in terms of age**.
- Although one of the benefits of randomization is that it may increase the likelihood of such comparability, it does not guarantee it. It is still possible that after we randomize, we may, by chance, find that most of the older patients are in one group and most of the younger patients are in the other.
- Our results would then be impossible to interpret because the higher-risk patients would be clustered in one group and the lower-risk patients in the other.
- Any difference in outcome between intervention groups may then be attributable to this difference in the age distributions of the two groups rather than to the effects of the intervention.

- In stratified randomization, we first stratify (stratum = layer) our study population by each variable that we consider important and then randomize participants to treatment groups within each stratum.
- Let us consider the example shown in Fig. 10.4. We are studying 1,000 patients and are concerned that sex and age are important determinants of prognosis.
- If we randomize, we do not know what the composition of the groups may be in terms of sex and age; therefore, we decide to use stratified randomization.
- We first stratify the 1,000 patients by sex into 600 males and 400 females. We then separately stratify the males by age and the females by age.
- We now have four groups (strata): younger males, older males, younger females, and older females.
- We now randomize within each group (stratum), and the result is a new treatment group and a current treatment group for each of the four groups.
- As in randomization without stratification, we end up with two intervention groups, but having initially stratified the groups, we increase the likelihood that the two groups will be comparable in terms of sex and age.

Stratified Randomization:



Data Collection on Subjects

- It is essential that the data collected for each of the study groups be of the same quality.
- We do not want any differences in results between the groups to be due to differences in the quality or completeness of the data that were collected in the study groups.

- Let us consider some of the variables about which data need to be obtained on the subjects:
  1. **TREATMENT (ASSIGNED AND RECEIVED)**
  2. **OUTCOME**
  3. **PROGNOSTIC PROFILE AT ENTRY**
  4. **MASKING (BLINDING)**

1. TREATMENT (ASSIGNED AND RECEIVED)

- **It is important to know**
  - Which treatment group the patient was assigned.
  - Which therapy the patient received.
    - It is important to know, if the patient was assigned to receive treatment A but did not comply.
    - A patient may agree to be randomized but may later change his or her mind and refuse to comply.
  - Whether a patient who was not assigned to receive treatment A may have taken treatment A on his or her own, often without the investigators knowing

Sheet# 12

- When you have two groups, it is important to know which group the patient is in and which treatment he received. Some patients will also be assigned to the placebo group but will take the treatment without the investigators' knowledge from outside resources. How can you tell they are compliant? Sometimes you can carry out urine and blood tests. You can also give the patients a list of medications and ask them to not take them.

2. OUTCOME

- The need for **comparable measurements** in all study groups is particularly true for **measurements of outcome**.
- Such measurements include both **improvement** (the desired effect) and **any side effects** that may appear.
- There is therefore a need for explicitly stated criteria for all outcomes to be measured in a study.
- Once the criteria are explicitly stated, we must be certain that they are measured comparably in all study groups.
- In particular, the potential pitfall of outcomes being measured more carefully in those receiving a new drug than in those receiving currently available therapy must be avoided.
- Blinding (masking), discussed later, can prevent much of this problem, but because blinding is not always possible, attention must be given to ensuring comparability of measurements and of data quality in all of the study groups.

Sheet# 13

- Any outcome should be stated clearly.
- It must also be measured comparably.
- You might pay more attention to the group with the intervention and thus, find more outcomes. Investigator masking can solve this. When can blinding be difficult? When the intervention is Surgical for example.

All-Cause Mortality Outcome ("Public Health Outcome")

- On occasion a medication or a preventive strategy for mortality that is effective with regard to the main outcome of interest does not increase event-free survival.
  - For example, in the 13-year follow-up of the European Randomized Study of Screening for Prostate Cancer, there was a reduction of approximately 27% in prostate cancer mortality.
  - However, overall mortality (also known as "public health outcome") was similar in the two study groups, thus suggesting that effectiveness of screening with regard to all-cause mortality was null.

Sheet# 14

- Mortality outcome is measured in two ways:

1) disease-specific:

For example: does the intervention (prostate screening) improve disease (prostate cancer) survival?

2) all cause: does the intervention (mostly in public health) improve overall survival?

- If the intervention improves both, it was worth being implemented.

PROGNOSTIC PROFILE AT ENTRY

- **If we know the risk factors for a bad outcome, we want to verify that randomization has provided reasonable similarity between the two groups in terms of these risk factors.**
  - For example, if age is a significant risk factor, we would want to know that randomization has resulted in groups that are comparable for age.
- **Data for prognostic factors should be obtained at the time of subject entry into the study, and then the two (or more) groups can be compared on these factors at baseline (i.e., before the treatment is provided).**
- **Another strategy to evaluate comparability is to examine an outcome totally unrelated to the treatment that is being evaluated.**
  - For example, if the randomized trial's objective is to evaluate a new medication for migraines, it is expected that mortality from cancer would be similar in the two groups.

Sheet# 15

- Based on whether a variable might be a risk factor, the study design needs to be changed to include or exclude the variable. See previous example: age and gender in Stratified randomization.
- How do you check if both groups are similar in qualities? Compare the mortality in both groups in an unrelated field. For example, if the intervention is migraine medication, check the cancer mortality. If it is similar, the groups are comparable. You do all this to decrease the number of factors that differentiate group A from group B.

MASKING (BLINDING)

- **Masking involves several components:**

1. **The subjects not to know which group they are assigned to.**
   - This is of particular importance when the outcome is a subjective measure, such as self-reported severity of headache or low back pain. If the patient knows that he or she is receiving a new therapy, enthusiasm and certain psychological factors on the part of the patient may operate to elicit a positive response even if the therapy itself had no positive biologic or clinical effect.

How can subjects be masked?

- One way is by using a **placebo**, an inert substance that looks, tastes, and smells like the active agent.
- However, use of a placebo does not automatically guarantee that the patients are masked (blinded).
  - Some participants may try to determine whether they are taking the placebo or active drug.
    - For example, in a randomized trial of vitamin C for the common cold, patients were blinded by use of a placebo and were then asked whether they knew or suspected which drug they were taking.

TABLE 10.4 Randomized Trial of Vitamin C and Placebo for the Common Cold: Results of a Questionnaire Study to Determine Whether Subjects Suspected Which Agent They Had Been Given

	SUSPECTED DRUG		
Actual Drug	**Vitamin C**	**Placebo**	**Total**
Vitamin C	40	12	52
Placebo	11	39	50
Total	51	51	102

$P < .001$.

From Karlowski TR, Chalmers TC, Frenkel LD, et al. Ascorbic acid for the common cold. A prophylactic and therapeutic trial. *JAMA*. 1975;231:1038. Copyright 1975, American Medical Association.

- As seen in Table 10.4, of the 52 people who were receiving vitamin C and were willing to make a guess, 40 stated they had been receiving vitamin C.
- Of the 50 who were receiving placebo, 39 said they were receiving placebo.
- **How did they know?**
  - They had bitten into the capsule and could tell by the bitter taste.
- **Does it make any difference that they knew?**
  - The data suggest that the rate of colds was higher in subjects who received vitamin C but thought they were receiving placebo than in subjects who received placebo but thought they were receiving vitamin C.
- Thus, we must be very concerned about lack of masking or blinding of the subjects and its potential effects on the results of the study, particularly when we are dealing with **subjective end points**.

Use of a placebo is also important for studying the rates of side effects and reactions.

- The Physicians' Health Study was a randomized trial of the use of aspirin to prevent myocardial infarctions. Table 10.5 shows the side effects that were reported in groups receiving aspirin and those receiving placebo in this study.
- Note the high rates of reported reactions in people receiving placebo. Thus it is not sufficient to say that 34% of the people receiving aspirin had gastrointestinal symptoms; what we really want to know is the extent to which the risk of side effects is increased in people taking aspirin compared with those not taking aspirin (i.e., those taking placebo).
- Thus the placebo plays a major role in identifying both the real benefits of an agent and its side effects.

TABLE 10.5 Physicians' Health Study: Side Effects According to Treatment Group

Side Effect	Aspirin Group (%)	Placebo Group (%)	*P*
GI symptoms (except ulcer)	34.8	34.2	.48
Upper GI tract ulcers	1.5	1.3	.08
Bleeding problems	27.0	20.4	<.00001

GI, Gastrointestinal.

Data from Steering Committee of the Physicians' Health Study Research Group. Final report on the aspirin component of the Ongoing Physicians' Health Study. *N Engl J Med*. 1989;321:129–135.

- Sometimes it is possible to use a medication in both the new therapy and in the placebo groups to prevent the occurrence of the most obvious side effects of the therapy.
  - In the aspirin example, a proton pump inhibitor, which is a class of medication that is used to prevent gastrointestinal symptoms from excess acid, could be given to both randomized groups, thus masking the participants with regard to the group to which they were allocated.
- Some patients knew they were taking aspirin when they developed GI symptoms. To create a masking effect, the investigators administered PPI to the patients to reduce the GI symptoms.

Blinding

1. Blinding the subjects

2. To mask (or blind) the observers or data collectors in regard to which group a patient is in (observer bias)

3. The masking of both participants and study personnel “double blinding.”

Other aspects of the design of randomized trials:

- Crossover
  - Planned
  - Unplanned
- Factorial Design

Planned crossover

- In this example, a new treatment is being compared with current treatment. Subjects are randomized to new treatment or current treatment (see Fig. 10.5A).
- After being observed for a certain period of time on one therapy and after any changes are measured (see Fig. 10.5B), the patients are switched to the other therapy (see Fig. 10.5C).
- Both groups are then again observed for a certain period of time (see Fig. 10.5D).
- Changes in group 1 patients while they are on the new treatment can be compared with changes in these patients while they are on the current treatment (see Fig. 10.5E).
- Changes in group 2 patients while they are on the new treatment can also be compared with changes in these patients while they are on the current treatment (see Fig. 10.5F).
- Thus, each patient can serve as his or her own control, holding constant the variation between individuals in many characteristics that could potentially affect a comparison of the effectiveness of two agents.

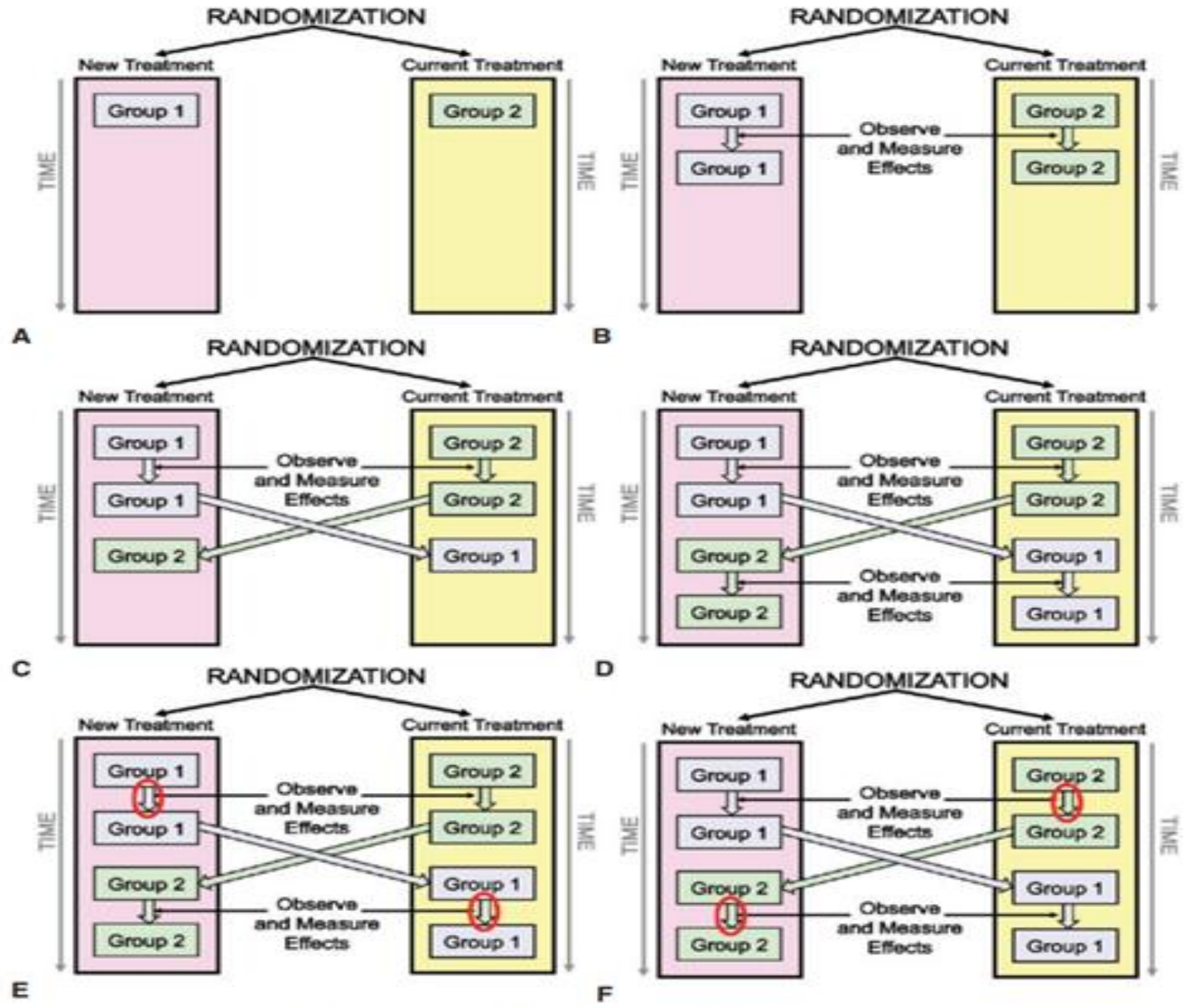


Fig. 10.5 (A–F) Design of a **planned crossover** trial. See discussion in text.

Cautions with planned crossover design

- **Carryover:**
  - For example, if a subject is changed from therapy A to therapy B and observed under each therapy, the observations under therapy B will be valid only if there is no residual carryover from therapy A.
  - There must be enough of a "**washout period**" to be sure none of therapy A, or its effects, remains before starting therapy B.
- **The order in which the therapies are given may elicit psychological responses**.
  - Patients may react differently to the first therapy given in a study as a result of the enthusiasm that is often accorded a new study; this enthusiasm may diminish over time.
  - We therefore want to be sure that any differences observed are indeed due to the agents being evaluated, and not to any effect of the order in which they were administered.
- **The planned crossover design is clearly not possible if the new therapy is surgical or if the new therapy cures the disease.**

Unplanned crossover

- Fig. 10.6A shows the design of a randomized trial of coronary bypass surgery, comparing it with medical care for coronary heart disease.
- Randomization is carried out after informed consent has been obtained.
- Although the initial design is straightforward, in reality, unplanned crossovers may occur.
- Some subjects randomized to bypass surgery may begin to have second thoughts and decide not to have the surgery (see Fig. 10.6B).
- They are therefore crossovers into the medical care group (see Fig. 10.6C).
- In addition, the condition of some subjects assigned to medical care may begin to deteriorate and urgent bypass surgery may be required (see Fig. 10.6B)—
- These subjects are crossovers from the medical to the surgical care group (see Fig. 10.6C).

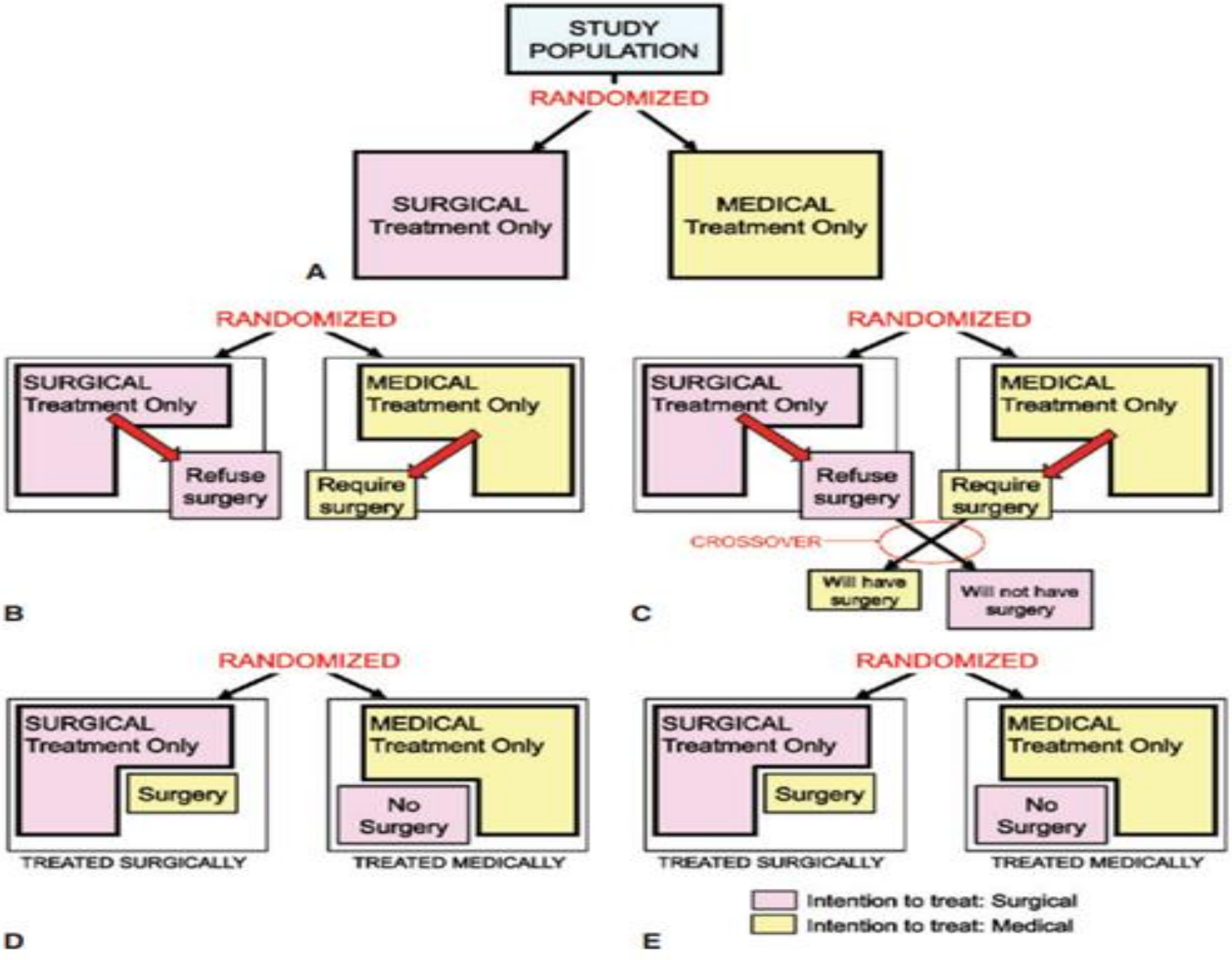


Fig. 10.6 (A–E) **Unplanned crossover** in a study of cardiac bypass surgery and the use of intention to treat analysis. (A) Original study design. (B–D) Unplanned crossovers. (E) Use of intention to treat analysis.

Sheet# 16

- After randomization, some participants wished to change groups. Some chose surgery while others chose medical treatment. If you follow their wishes, the study will be biased. So, you must continue using the first groups you started with but move the participants according to the treatment they wish. More details on page 47.

- The patients seen on the left in Fig. 10.6D are now treated surgically, and those on the right in this figure are treated medically.
- Those treated surgically include some who were randomized to surgery (shown in pink) and some who crossed over to surgery (shown in yellow).
- Those treated medically include some who were randomized to medical treatment (shown in yellow) and some who crossed over to medical treatment (shown in pink).

Intention to treat analysis and as treated analysis

- Unplanned crossovers pose a serious challenge in analyzing the data. If we analyze according to the original assignment (called an intention to treat analysis), we will include in the surgical group some patients who received only medical care, and we will include in the medical group some patients who had surgery.
- In other words, we would compare the patients according to the treatment to which they were originally randomized, regardless of what treatment actually occurred.
- Fig. 10.6E shows an intention to treat analysis in which we compare the group in pink (randomized to surgical treatment) with the group in yellow (randomized to medical treatment).
- If, however, we analyze according to the treatment that the patients actually receive (as treated analysis), we will have broken, and therefore lost the benefits of, the randomization.

- No perfect solution is available for this dilemma.
- Current practice is to perform the primary analysis by intention to treat—according to the original randomized assignment.
- The bottom line is that because there are no perfect solutions, the number of unplanned crossovers must be kept to a minimum.
- Obviously, if we analyze according to the original randomization and there have been many crossovers, the interpretation of the study results will be questionable.

Factorial Design

- Assuming that two drugs are to be tested, the anticipated outcomes for the two drugs are different, and their modes of action are independent, one can economically use the same study population for testing both drugs.

- This factorial type of design is shown in Fig. 10.7.

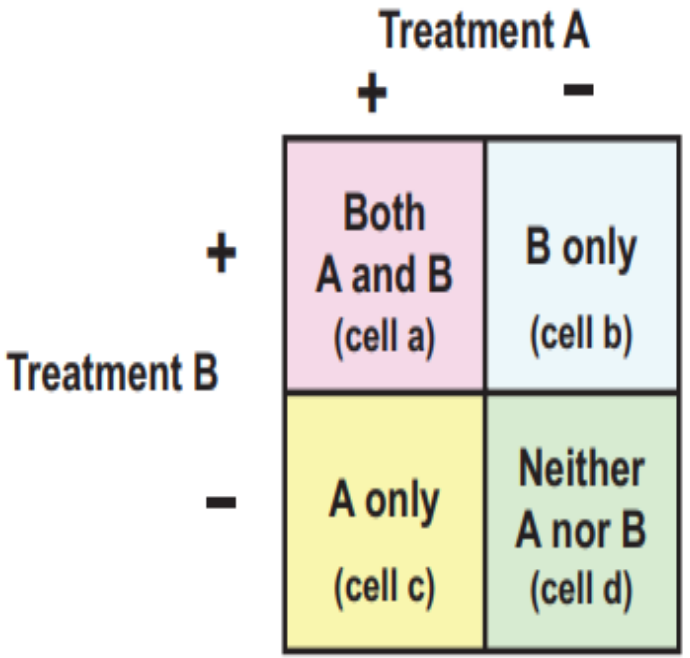


Fig. 10.7 Factorial design for studying the effects of two treatments.

Sheet# 17

- When do you use factorial design? When you have two medications to test, and both are independent and do not influence each other. Instead of getting controls for both medications, you only use both medications. Medication A + B
- Medication A alone
- Medication B alone
- Neither medication (no treatment group = comparison group)
- = less costs in this design

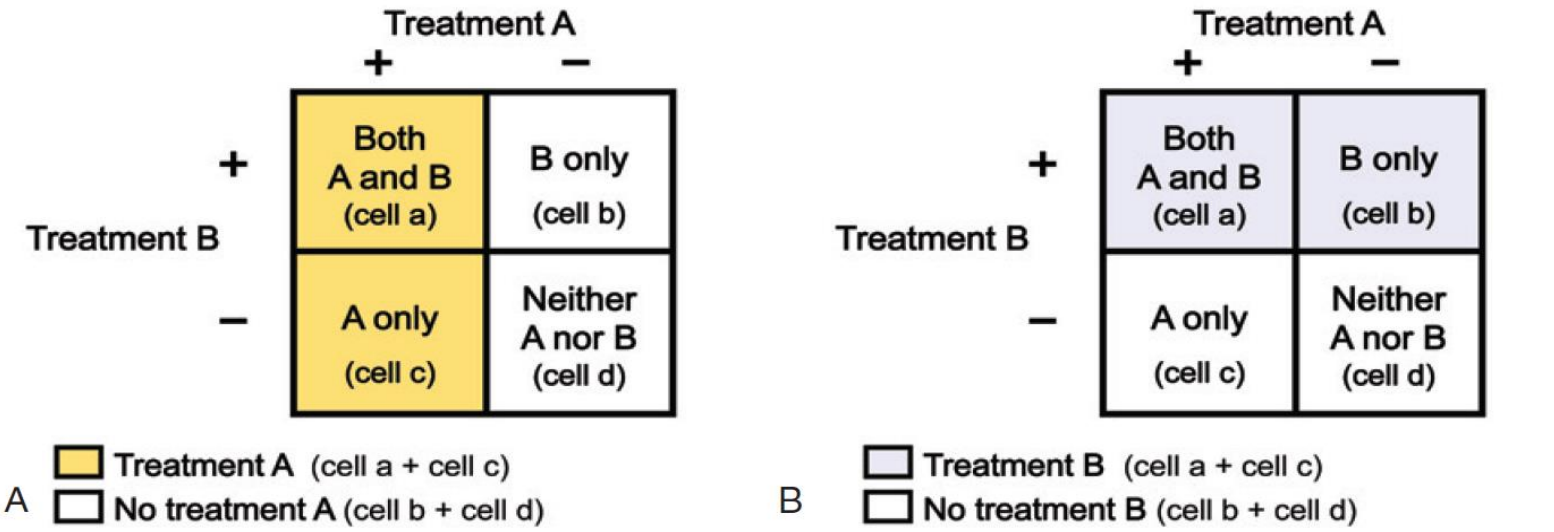


Fig. 10.8 (A and B) Factorial design. (A) The effects of treatment A *(orange cells)* versus no treatment A. (B) The effects of treatment B *(purple cells)* versus no treatment B.

- If the effects of the two treatments are indeed completely independent, we could evaluate the effects of treatment A by comparing the results in cells a + c to the results in cells b + d (Fig. 10.8A).
- Similarly, the results for treatment B could be evaluated by comparing the effects in cells a + b to those in cells c + d (see Fig. 10.8B).
- In the event that it is decided to terminate the study of treatment A, this design permits continuing the study to determine the effects of treatment B.

Example – Factorial design

- An example of a factorial design is seen in the Physicians' Health Study.16 More than 22,000 physicians were randomized using a 2 × 2 factorial design that tested aspirin for primary prevention of cardiovascular disease and beta carotene for primary prevention of cancer.
- Each physician received one of four possible interventions: both aspirin and beta carotene, neither aspirin nor beta carotene, aspirin and beta carotene placebo, or beta carotene and aspirin placebo.
- The resulting four groups are shown in Figs. 10.9 and 10.10.
- The aspirin part of the study (Fig. 10.11A) was terminated early, on the advice of the external data monitoring board, because a statistically significant 44% decrease in the risk of first myocardial infarction was observedin the group taking aspirin.
- The randomized beta carotene component (see Fig. 10.11B) continued until the originally scheduled date of completion.
- After 12 years of beta carotene supplementation, no benefit or harm was observed in terms of the incidence of cancer or heart disease or death from all causes. Subsequent reports have shown greater risk of cancer with beta carotene in smokers

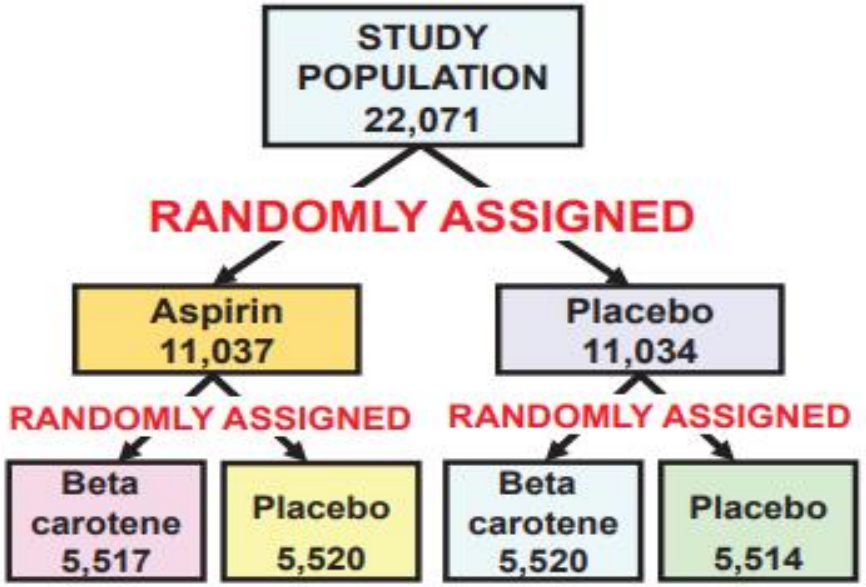


Fig. 10.9 Factorial design used in a study of aspirin and beta carotene.

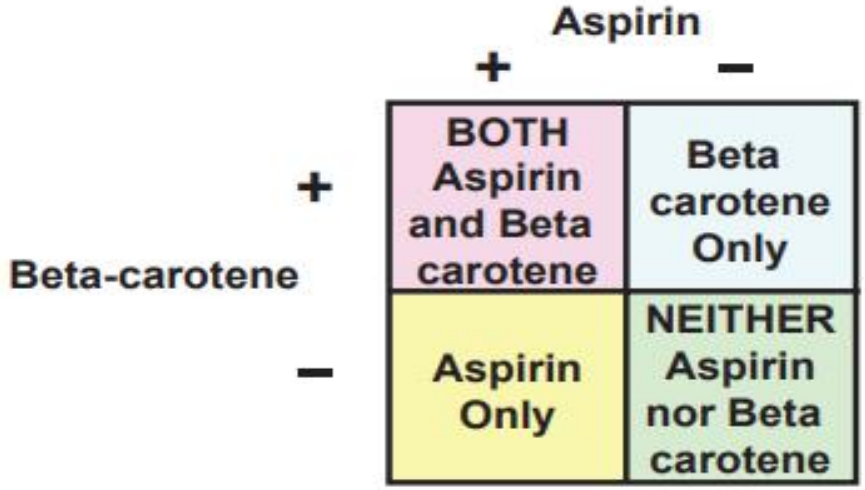
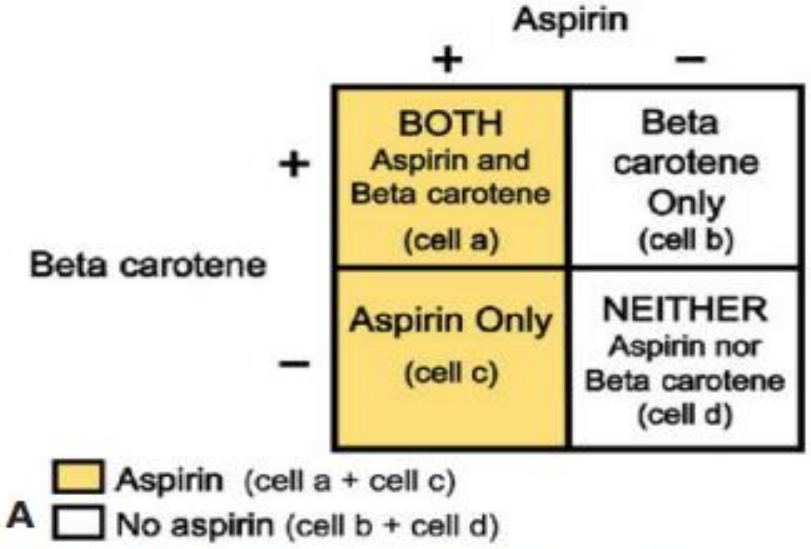


Fig. 10.10 Factorial design of the study of aspirin and beta carotene in 2 × 2 table format.



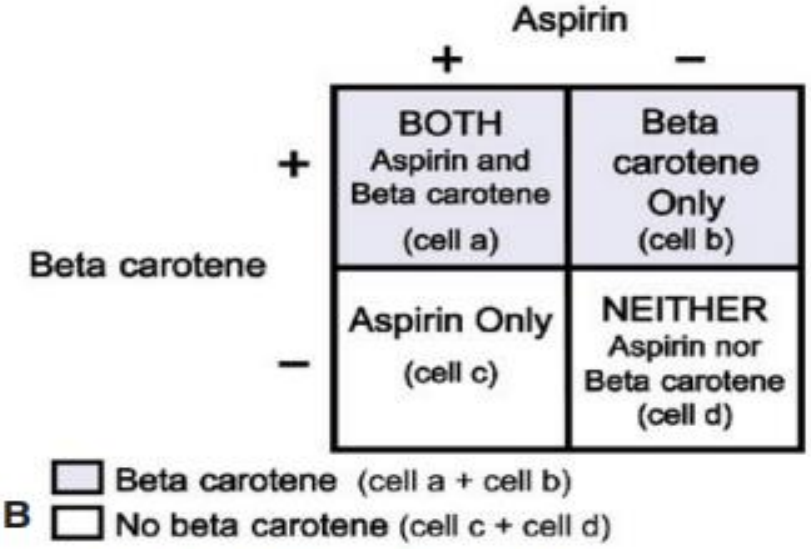


Fig. 10.11 (A and B) Factorial design. (A) The effects of aspirin *(orange cells)* versus no aspirin. (B) The effects of beta carotene *(purple cells)* versus no beta carotene.

Noncompliance

- Patients may agree to be randomized but following randomization they may not comply with the assigned treatment.
- **Noncompliance may be overt or covert:**
  - On the one hand, people may overtly articulate their refusal to comply or may stop participating in the study.
    - These noncompliers are also called **dropouts** from the study.
  - On the other hand, people may just stop taking the agent assigned without admitting this to the investigator or the study staff.
- Whenever possible, checks on potential noncompliance are built into the study.
  - These may include, for example, urine tests for the agent being tested or for one of its metabolites.
- Biggest problem with randomized trials is noncompliance of the participants. They are called drop-outs.

Drop-ins

- Another problem in randomized trials has been called drop-ins.
- Patients in one group may inadvertently take the agent assigned to the other group.
  - For example, in a trial of the effect of aspirin for prevention of myocardial infarction, patients were randomized to aspirin or to no aspirin.
  - However, a problem arose in that, because of the large number of over-the-counter preparations that contain aspirin, many of the control patients might well be taking aspirin without knowing it.
  - Two steps were taken to address this problem:
    1. controls were provided with lists of aspirin-containing over-the-counter preparations that they should avoid
    2. urine tests for salicylates were carried out both in the aspirin group and in the controls.

Sheet# 18

- drop-ins who take the treatment without our knowledge. Explained previously. Perform tests to determine if participant is a drop-in.
- Sometimes dropouts = drop-ins. If they are equal, they cancel the effect of one another.

- The net effect of noncompliance on the study results will be to reduce any observed differences (i.e., driving the difference toward the null) because the treatment group will include some who did not receive the therapy, and the no-treatment group may include some who received the treatment.
- Thus the groups will be less different in terms of therapy than they would have been had there been no noncompliance, so that even if there is a difference in the effects of the treatments, it will appear much smaller.

MAINTENANCE OF COMPLIANCE

- Selecting high risk people as participants in study population.
- Frequent contacts with the participants through phone calls, home visits, clinic visits.
- Providing calendar packs to the participants and asking them to stick on to calendar packs without fail.
- Giving incentives like free medical aid in future, giving some gifts.

NON-COMPLIANCE

- Non-compliance decreases the statistical power of the trial which speaks about the validity (truth of the results)
- Extent of non-compliance is directly proportional to the duration and complexity of the trial.
- Compliance is difficult when the end –points are time taking like incidence of cancers or death