COURSE 5: RANDOMIZATION & BLINDING

OVERVIEW OF RANDOMIZATION

Randomization: The process by which each subject has the same chance of being assigned to either intervention or control.

Neither the subject **nor** the investigator should know the treatment assignment before the subject's decision to enter the study!!!!
$\Rightarrow$ This removes investigator bias.

Bias may be defined as systematic error, or "difference between the true value and that actually obtained due to all causes other than sampling variability." Bias is the "killer" of the study. Therefore, bias reduction is an extremely important issue in the trial design phase, as well as in the trial implementation phase.

Randomization:
- Tends to produce groups that are comparable w.r.t. known or unknown risk factors
- Guarantees the validity of statistical tests

Methods of randomization:
- Phone call to central office
- Sealed envelopes

Check eligibility **before** randomization.

Randomize as close to treatment time as possible (to avoid death or withdrawal before treatment start).

There are several methods for making random treatment assignments. Many attempt to **balance** treatment groups over time, over stratification factors, or both. In the following, we assume **equal** allocation of patients to each treatment (i.e., 1:1 randomization).

BASIC RANDOMIZATION METHODS

Simple Randomization

Methods

- Toss a coin:
    - o H → intervention
    - o T → control

- Generate a random digit (Table 5.2, Pocock, p. 74 [use a random starting point] or use a calculator or computer program):
    - o Even # → intervention, odd # → control
            or
    - o 0 to 4 → intervention, 5 to 9→ control

- Do **not** use alternating assignment (ABAB....)
    - o No random component; investigator knows next assignment.

Table 5.2. A table of random numbers

0 5	2 7	8 4	3 7	4 1	6 8	3 8	5 1	5 6	9 6	8 1	8 0	4 7	8 8	7 4
5 9	7 2	4 0	2 3	6 3	1 8	5 0	2 6	0 9	9 6	9 2	1 8	8 5	0 3	7 9
2 5	9 8	4 3	8 9	5 2	8 4	6 4	4 2	7 5	4 4	9 2	8 1	8 6	9 3	2 2
3 0	3 6	6 5	1 3	1 8	7 7	6 0	0 9	4 3	6 8	0 5	9 5	1 7	5 2	4 2
6 8	1 5	6 7	7 5	7 5	3 4	8 8	0 8	8 8	6 5	2 1	2 8	1 2	8 9	5 5
7 3	6 5	8 3	7 8	1 0	7 9	7 5	5 9	9 9	7 3	8 9	8 3	8 4	5 8	3 2
7 0	1 0	0 7	5 4	1 3	0 1	9 6	6 9	9 5	5 ‚3	2 5	6 4	4 6	1 6	5 6
2 8	1 5	3 3	1 2	8 3	9 0	2 0	2 5	1 5	5 7	0 0	0 4	4 7	4 7	4 4
1 5	6 6	7 2	9 3	0 9	2 4	6 1	1 7	9 1	4 0	6 0	2 5	1 5	2 8	3 3
0 0	2 1	8 2	3 2	0 0	6 8	3 6	5 9	7 3	4 5	3 4	5 0	0 5	6 6	9 6
9 9	6 7	3 8	3 2	5 4	8 3	2 5	3 3	9 5	6 3	5 2	2 1	4 9	5 0	3 4
3 6	8 0	5 3	3 8	1 5	2 1	2 1	4 8	3 0	5 3	7 8	9 1	3 4	7 7	2 7
2 9	1 9	2 9	1 8	8 4	0 9	9 0	8 7	0 4	9 3	3 9	0 3	9 0	4 1	7 9
5 4	6 3	6 0	3 1	6 5	3 5	9 5	4 5	1 4	0 8	7 9	6 8	0 4	9 9	9 9
0 2	2 7	2 4	6 0	5 7	0 5	1 6	4 7	5 0	2 2	0 1	2 6	4 9	5 4	9 3
7 3	2 5	6 8	3 2	4 4	6 1	2 5	9 7	9 6	2 2	9 9	9 1	4 2	5 0	4 3
9 0	2 8	6 7	2 6	9 7	6 3	0 4	2 7	5 6	2 4	8 0	1 1	0 1	6 3	3 4
6 1	3 3	9 3	6 2	4 0	6 0	6 8	9 9	1 9	0 2	0 3	3 2	8 7	7 4	6 4
7 0	7 7	1 1	5 7	2 1	3 4	4 6	1 1	2 4	7 2	4 2	4 4	9 2	1 6	5 9
8 3	6 4	0 7	2 2	9 3	3 2	5 5	9 1	5 3	6 3	7 8	0 7	7 5	8 7	1 2
7 3	3 2	5 8	8 6	2 5	0 4	6 6	0 0	5 8	3 9	0 1	5 8	3 0	9 3	0 6
7 6	6 3	0 0	8 8	7 4	2 3	9 5	8 1	2 6	7 1	4 8	6 9	7 1	9 6	9 9
4 3	2 8	9 0	4 5	7 9	0 5	2 3	5 8	7 9	1 6	4 2	7 5	6 2	4 3	8 4
0 5	6 6	0 1	5 1	7 2	8 7	3 5	1 6	8 9	9 6	2 4	5 5	0 9	2 5	6 6
0 6	1 4	3 7	7 4	1 2	8 9	1 6	2 7	7 5	2 0	4 0	4 8	3 6	3 9	2 7
6 9	7 7	8 5	1 4	9 6	1 8	4 0	6 3	4 3	8 3	5 2	2 5	9 5	5 7	6 8
2 1	4 8	6 3	9 7	7 9	2 0	3 1	2 1	5 6	0 3	1 9	2 3	4 3	8 4	6 2
4 7	6 9	5 6	1 7	5 8	6 8	2 9	2 5	9 5	3 1	0 9	8 9	6 8	6 0	5 9
2 5	5 7	7 5	8 3	3 1	9 1	4 3	7 9	7 7	0 3	5 5	8 8	7 5	6 9	2 0
5 6	5 2	8 8	7 4	5 1	6 3	2 4	0 1	6 5	6 9	2 0	7 3	8 6	7 3	2 1
4 8	3 6	7 3	7 5	3 9	5 9	9 8	2 5	7 2	5 8	9 4	1 3	1 7	2 7	5 5

Examples

<u>Two treatments</u>
0 to 4 → A
5 to 9 → B

Random digits:	7	2	4	0	2	3	6	3	1	8	...
Treatment assigned:	B	A	A	A	A	A	B	A	A	B	...

Number assigned to each group: 7A, 3B

<u>Three treatments</u>
1 to 3 → A
4 to 6 → B
7 to 9 → C
(0 → ignore)

Random digits:	7	2	4	0	2	3	6	3	1	8	...
Treatment assigned:	C	A	B	-	A	A	B	A	A	C	...

(5A, 2B, 2C)

Number assigned to each group: 5A, 2B, 2C

Pros and Cons

Pros
- Easy to implement

Cons
- At any point in time, there may be an imbalance in the number of subjects on each treatment.

Balance improves as n increases
- With n = 20, the chance of a 12:8 split or worse is ≈ **50%**.
- With n = 100, the chance of a 60:40 split or worse is still > **5%**.

⇒ It is desirable to restrict randomization to ensure similar treatment numbers throughout the trial.

Three approaches to restricting randomization will be given: replacement, permuted block, and biased coin.

Replacement Randomization

For a planned trial with 2n patients, specify an amount of imbalance that would be unacceptable. For example, with n = 100 patients on each treatment, the treatment imbalance (# on treatment A – # on treatment B) should be less than 6 at any point in time. Using simple randomization, generate a randomization list. If the imbalance is unacceptable using the pre-specified criterion, then generate a new list. Repeat, if necessary, until an acceptable list is obtained.

Random Permuted Blocks

Blocking is used to equalize the number of subjects on each treatment. For each block of patients of size b, b/2 subjects are assigned to each treatment.

Example

With blocks of size 4, after every 4 patients there will be an equal number assigned to each treatment.

Block 1	Block 2	Block 3	etc.
ABAB	BAAB	BABA	

Method 1

Write down all permutations for the given block size, b. For block size = b, there are

$$\binom{b}{b/2} = \frac{b!}{(b/2)!(b/2)!}$$

arrangements.

For example, for b = 4, there are

$$\binom{4}{2} = \frac{4!}{2!2!} = 6$$

arrangements: AABB, ABAB, BAAB, BABA, BBAA, and ABBA. For each block, randomly choose one of these arrangements.

Method 2

In each block, generate a random number for each treatment, then rank the treatments in order.

Example

Treatments in any order	Random #	Rank	Treatments in rank order
A	07	1	A
A	73	3	B
B	87	4	A
B	31	2	B

Choose a new set of random numbers for each block.

Considerations

With blocked randomization, where b = block size, the number in each group never differs by more than b/2.
⇒ This ensures treatment balance through whole accrual period.

Blocking factor, b, should **not** be known to investigators (if known, the last treatment in each block is predictable).

A trial without further stratification should have a fairly large block size (say, b = 10 to 20) to reduce predictability.

Do **not** use blocks of size 2.

Block size can be varied over time, even **randomly**.

Biased Coin

If the number of subjects already on each treatment (n_1 and n_2) is equal ($n_1 \approx n_2$), then randomize to either treatment with P = 1/2.

If $n_1 > n_2 + C$, then increase P(treatment 2) to be > 1/2.
If $n_2 > n_1 + C$, then increase P(treatment 1) to be > 1/2.

(C = unacceptable level of imbalance between group sizes)

Larger P $\Rightarrow$ imbalance will be corrected quickly.
• Suggested P $\approx$ 2/3

Pros
• Next assignment cannot be predicted.
• Statistical power is greater with equal allocation.

RANDOMIZATION METHODS FOR TREATMENT BALANCE OVER PROGNOSTIC FACTORS & INSTITUTION

Stratified Permuted Block Randomization

Simple randomization guarantees that treatment balance within prognostic factors will occur **on average**. However, in a **particular** study, especially with small trials, the imbalance may be great.

Stratified randomization guarantees treatment balance within prognostic factors.

> ⇒ Overall treatment comparison is unconfounded with other factors. (Imbalances can be adjusted using regression analysis, but if conclusions are different, doubt is cast on the results.)

Very large trials (say, > 500 patients) may not require stratification.

Process:
1. First define strata.
2. Randomization is performed **within each stratum**, and is usually blocked. Use different blocking patterns in each cell.

Example

$$\text{Age} \begin{Bmatrix} < 40 \\ 41-60 \\ 61+ \end{Bmatrix} \text{ and gender} \begin{Bmatrix} M \\ F \end{Bmatrix}$$

Total number of strata = 3 x 2 = 6

	Male	Female
Age < 40	ABBA, BAAB, ...	BABA, BAAB, ...
Age 41 to 60	BBAA, ABAB, ...	ABAB, BBAA, ...
Age ≥ 60	AABB, ABBA, ...	BAAB, ABAB, ...

The block size should be small (b = 2 or 4) to maintain balance in small strata, and to ensure that the overall imbalance is not too great. With several strata, predictability should not be a problem.

Increased number of stratification variables or increased number of levels within strata ⇒ fewer patients per stratum. In small samples, sparse data in many cells defeats the purpose of stratification.

66

Note: Stratification factors also should be used in analysis. Otherwise, the overall test will be conservative.

Minimization

- Balances treatments simultaneously over several prognostic factors (strata)
- Does **not** balance **within** cross-classified stratum cells; balances over the marginal totals of each stratum separately
- Is used when the number of stratum cells is large relative to sample size (stratified design would yield sparse cells)
- Can be computerized

Method

Keep a current list of the total patients on each treatment for each stratification factor level.

Example 1

Three stratification factors:
- Gender (2 levels)
- Age (3 levels)
- Disease stage (3 levels)

After 50 patients, the list is as follows:

		Treatment A	Treatment B
Gender:	Male	16	14
	Female	10	10
Age:	≤ 40	13	12
	41 to 60	9	6
	≥ 61	4	6
Disease stage:	Stage I	6	4
	Stage II	13	16
	Stage III	7	4
Total		**26**	**24**

Say the 51st patient enrolled in the study is male, age ≥ 61, stage III. Consider the lines from the table above for that patient's stratification levels only:

	Treatment A	Treatment B	Sign of difference
Male	16	14	+
Age ≥ 61	4	6	-
Stage III	7	4	+
Total	**27**	**24**	**2 +, 1 -**

Two possible criteria:
1. Count only the direction (sign) of the difference in each category. A is ahead in two categories out of three.
    ⇒   Assign next patient to B.

2. Add the total overall categories (27A vs. 24B). A is ahead.
    ⇒   Assign next patient to B.

These two criteria will usually agree, but not always.

Choose one of the two criteria to be used for the entire study. Both criteria will lead to reasonable balance.

Either criterion may lead to a "tie."

Example 2
If the 51st patient is female, age ≥ 61, stage I:

	Treatment A	Treatment B	Sign of difference
Female	10	10	0
Age ≥ 61	4	6	-
Stage I	6	4	+
Total	**20**	**20**	**1 +, 1 -**

Here, both criteria lead to a tie. In this case, simply randomize with P(treatment A) = P(treatment B) = 1/2.

Although the procedure is actually deterministic, it is not predictable without knowing the entire patient list. In a multi-center trial, such a list would not be available to individual centers. In a single-center setting, some randomness could be added to the method: If the next patient should be assigned to treatment A by one of the criteria, then instead of assigning to A automatically, assign to A with probability, say, 3/4, and to B with probability 1/4.

Generalizations of the method

- More than two treatments:
    - o **Criterion 1** (using signs of differences only) is **not** easily generalized.
    - o **Criterion 2** (using sums for each treatment) extends easily to this case by adding columns to the table for each additional treatment. The treatment with the smallest sum is then assigned to the next patient.

- Weighting stratification factors differently:
    - o **Criterion 1**: Instead of just using the signs (+,-), multiply the signs by weights and add over stratification.
    - o **Criterion 2**: Multiply the stratification level totals by the weights before adding.

Example 3

Weights:
- Gender = 1
- Age = 1
- Stage = 2

Assume the 51st patient is male, age $\geq$ 61, stage III:

	Weight	Treatment A	Treatment B	Weighted sign of difference
Male	1	16	14	+1
Age $\geq$ 61	1	4	6	-1
Stage III	2	7 x 2 = 14	4 x 2 = 8	+2
Total		**34**	**28**	**+2**

Criterion 1: The signs are +, -, +, so the weighted signs are +1, -1, +2, and the sum = +2. A positive sum implies A is ahead.

⇒   Assign the patient to B.

Criterion 2: The totals for stage are multiplied by 2. The lead for A is even larger than before.

⇒   Assign the patient to B.

In this example, the weighted and unweighted treatment assignments are the same, but this will not always be the case.

The method of minimization requires a new calculation for each new patient assignment (the randomization list cannot be prepared ahead of time). If done by hand, the calculations are facilitated by index cards, one card for each stratification level, which are updated for each new patient.

Alternatively, the algorithm is easy to program for a computer. If the computer is down, simple randomization can be used.

Minimization is an excellent method for achieving balance in a relatively small study with several stratification factors.

Careful: Balance by margins does <u>not</u> guarantee overall treatment balance, or balance within stratum cells.

Stratifying by Institution

Multi-center trials often have, say, 10 to 50 participating institutions.

Using the minimization technique, adding institution as a stratification factor is no problem.

Using random permuted blocks within strata, adding institution as a stratification factor will probably lead to sparse cells (and potentially more cells than patients!). Consider simply adding a criterion that if the institution imbalance is greater than, say, 4, then assign the next patient to the lagging treatment (with high probability).

OTHER RANDOMIZATION METHODS

Pre-randomization

A method used when there may be resistance among physicians to randomizing patients.

Method

When an eligible patient is identified, **first** determine the random treatment assignment, **then** approach the patient to ask for consent to participate in the study. (May be easier to discuss the study with the patient if the treatment they would receive is known).

All patients are followed after randomization. For analysis, patients are grouped **as randomized**, ignoring the treatment actually given. (This analysis is not biased by physician or patient selection biases, but may *under*estimate the actual treatment effects. Patient refusing the randomized treatment and receiving the alternate treatment will dilute the observed treatment effect.)

For a pre-randomized trial to be successful, the proportion of patients refusing the randomized treatment must be very small (less than 10%). (A well-run NSABP trial accomplished this, but subsequent trials did not, and were difficult to interpret.)

Additional sample size is required to compensate for those refusing. Pre-randomization is sometimes implemented to boost accrual, but failure to treat patients as assigned leads to the need for even more patients.

Response-adaptive Randomization

These methods "randomize" patients based on the response of previous subjects.

These designs are controversial and not commonly used.

Play-the-winner (PW) design

This design can be used when patient response is determined quickly. In practice, PW randomization is used if the investigator feels fairly strongly that the new treatment is more effective than conventional therapy before the study is started.

For the first subject, toss a coin:
- H→ treatment A
- T → treatment B

If response = success (S), then second subject receives the same treatment. Stay with the winner until a failure (F) is observed. At failure, assign next subject to other treatment. For example:

Patient #	1	2	3	4	5	6	7	8	...
Treatment A	S	F				S	F		
Treatment B			S	S	F			S	...

Pros
- More patients receive the better treatment.

Cons
- Investigator knows the next assignment.
- May lead to loss of statistical power if final sample sizes are quite unequal.

Two-armed bandit

Treatment assignment probabilities depend on observed success probabilities at each time point.

Pros
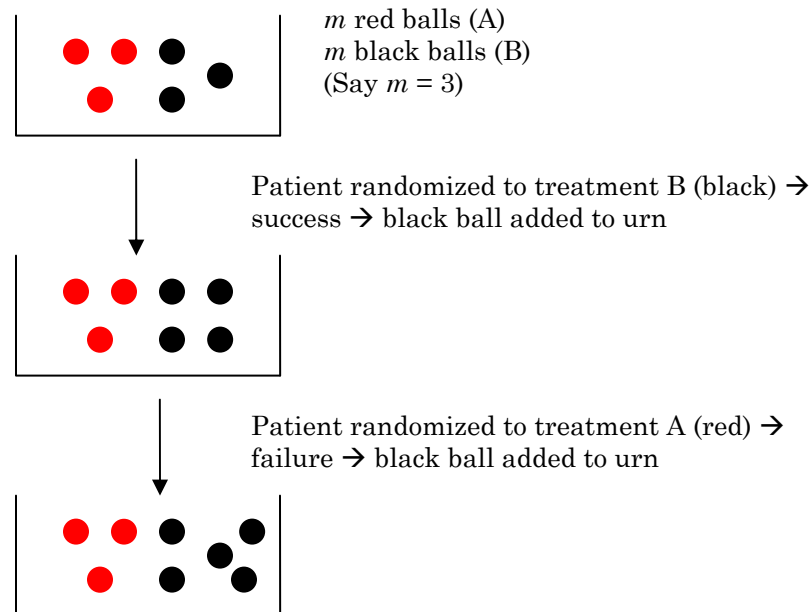- Attempts to maximize the number of subjects on the "superior" treatment.

Cons
- When unequal treatment numbers result, there is loss of statistical power in the treatment comparison.

Biased urn model

The biased urn model is one method for adapting the probability of treatment assignment based on success/failure in previous study subjects. The method is as follows:

1. Suppose that randomization to one of two treatments is made by the logical equivalent of drawing balls labeled A or B from an urn, with replacement.
2. For the first patient, the urn contains m balls of each type (A and B). The first patient is assigned on the basis of a random draw.
3. If the assigned treatment "fails," a ball of the other type is added to the urn. The next patient therefore has a higher probability of receiving the other treatment.

4. If the assigned treatment "succeeds," a ball of the same type is added to the urn. Thus, the next patient has a higher probability of receiving the same treatment.

m red balls (A)
m black balls (B)
(Say *m* = 3)

Patient randomized to treatment B (black) →
success → black ball added to urn

Patient randomized to treatment A (red) →
failure → black ball added to urn

Example of response-adaptive randomization

Extracorporeal circulation in neonatal respiratory failure

Bartlett et al. (Pediatrics 1985;76:479-487) implemented a trial of extracorporeal circulation with a modified heart-lung machine (extracorporeal membrane oxygenation [ECMO]) in infants with severe respiratory failure.
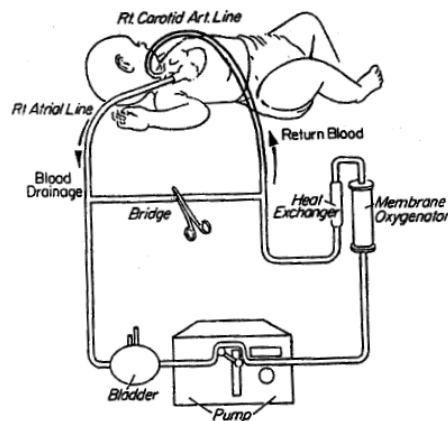
A play-the-winner design (using a biased urn) was chosen because:
- Outcome of treatment in these patients is known quickly.
- Most ECMO patients were expected to survive, while most control (conventional treatment) patients were expected to die. Therefore, significance could be achieved with a small overall sample size.
- PW design addressed both scientific and ethical demands. Scientifically, investigators felt obligated to perform a randomized trial. Ethically, investigators did not want to withhold lifesaving treatment from study subjects simply for the sake of conventional randomization.

Eligibility for randomization:
- Newborns > 2 kg
- Severe respiratory failure with 80% to 100% risk of dying based on:
    - Acute deterioration
    - Unresponsiveness
    - Barotrauma
    - Diaphragmatic hernia
    - High newborn pulmonary insufficiency index
- No contraindications

Randomization scheme:
1. The first patient would be randomized to ECMO or conventional therapy with equal probability (one of each ball in the urn).
2. For each patient who survived on ECMO or died on conventional therapy, one ECMO ball would be added to the urn.
3. For each patient who survived on conventional therapy or died on ECMO, one conventional therapy ball would be added to the urn.
4. Randomization would continue until 10 of one ball type had been added to the urn. Then randomization would cease, and all patients would be assigned to the successful therapy.

Results:
- Patient 1 randomized to ECMO → survived
    - → Subsequent odds of randomization to ECMO = 2:1
- Patient 2 randomized to conventional therapy → died
    - → Subsequent odds of randomization to ECMO = 3:1
- Patient 3 randomized to ECMO → survived
    - → Subsequent odds of randomization to ECMO = 4:1
- All remaining patients randomized to ECMO → survived
- Final results:
    - 11 ECMO patients → all survived
   - 1 control patient → died

Unequal Randomization

Reasons for unequal randomization:
- To gain greater experience using a new treatment.
- To improve accrual if enthusiasm for new treatment is high.

Equal allocation of patients to treatments gives maximum statistical power for the treatment comparison. How much power is lost when allocation is unequal?

(In the following example, the power with equal allocation is 0.95).
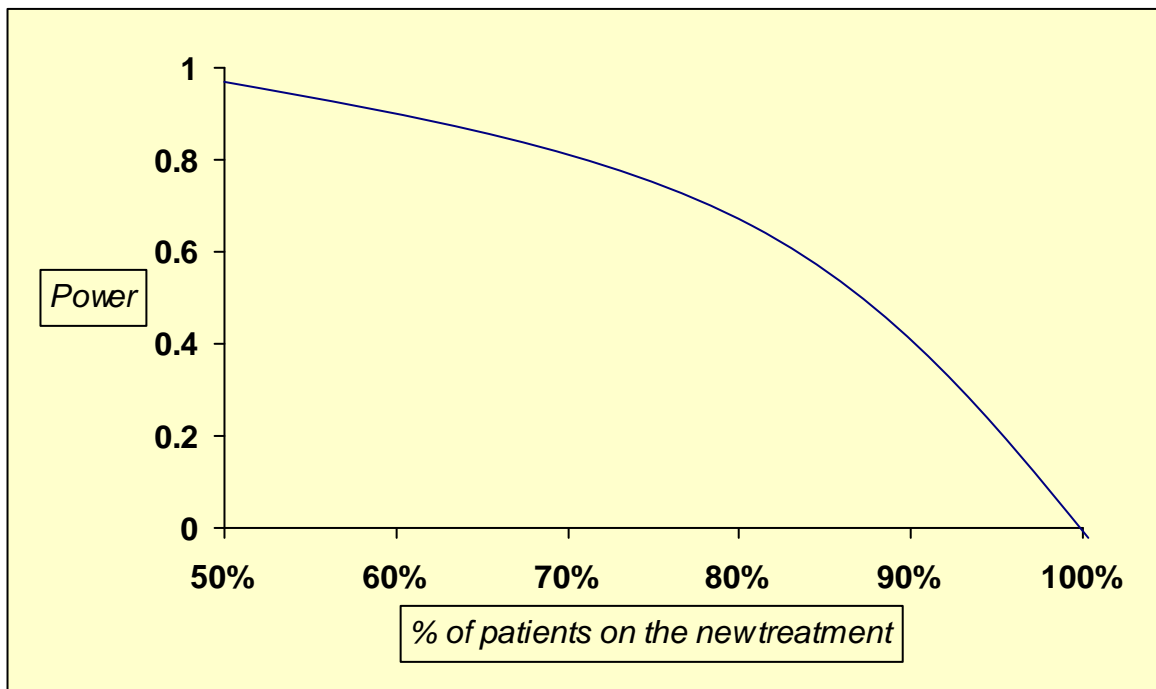
Reduction in power of a trial as the proportion on the new treatment is increased

Randomizing in a 2:1 ratio (2/3 of patients on new treatment):
> ⟹ Power decreases from 0.95 to 0.925 (not much loss!).

More extreme ratios lead to greater power losses. For example, a ratio of 4:1 (4/5 of patients on new treatment):
> ⟹ Power decreases from 0.95 to 0.82.

Moderately unequal randomization is statistically feasible, and may be especially useful in phase II randomized trials.

RANDOMIZATION METHOD RECOMMENDATIONS

1. **Large study** (several hundred participants)

    One center →   Blocked randomization

    Multicenter →   Stratified (by center)
                            Blocked

2. **Small study** ($n \approx 100$)

    One center →   Stratified (by 1 or 2 important risk factors)
                            Blocked

    Multicenter →   Stratified (by center + risk factors)
                            Blocked

3. **Very small study** ($n \approx 50$)

    One center →   Adaptive minimization

    ($n \approx 100$, many risk factors need to be balanced)

RANDOMIZATION EXAMPLE

Patients with stable lung function were eligible to enter the hospital for instruction in negative pressure ventilation (NPV) and to undergo baseline testing.

Randomization to active or sham NPV took place on the second day of hospitalization after a final evaluation designed to eliminate patients who were unable to perform the required tests and training or in whom testing seemed contraindicated. Eligible patients were randomized according to a stratified block randomization scheme. Stratification was based on whether or not the patient had been receiving home oxygen; home oxygen status was considered a composite measure of many different baseline factors and as such was an important variable upon which to achieve treatment balance.

Randomization proceeded within strata according to a permuted block scheme with a block size, or balancing interval, varying randomly between 8 or 12 according to the outcome of a computer generated random number. This ensured that the cumulative number of assignments to each treatment would be in balance after each block of assignments had been made. Treatment assignments were obtained by telephone request to the trial secretary who was located away from the site of clinical activities.

Shapiro SH, et al. A randomized clinical trial of negative pressure ventilation in severe chronic obstructive pulmonary disease: design and methods. *J Clin Epidemiol* 1991;44(6):483-496.

SUMMARY OF RANDOMIZATION METHODS

1.   **Simple randomization**

Toss a coin
Random digit

2.   **Random permuted block**

Block size/2   =   Maximum difference between group sizes

Block size   =   4 $\rightarrow$ 6 arrangements
Block size   =   6 $\rightarrow$ 20 arrangements
Block size   =   20 $\rightarrow$??? arrangements

3.   **Biased coin**

C, P $\approx$ 2/3

4.   **Stratified permuted block randomization**

Stratify by:
- Main risk factor(s)
- Medical center

5.   **Minimization**

Marginal total
+, -
Can be weighted
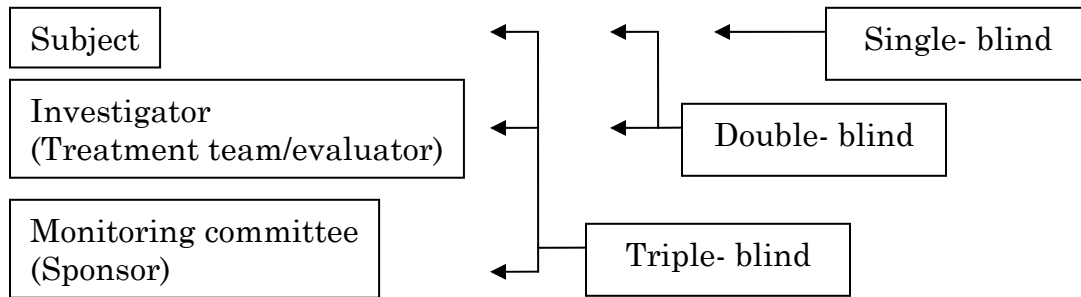
6.   **Pre-randomization**

7.   **Response-adaptive randomization**

8.   **Unequal randomization**

BLINDED STUDIES

Blinding

Keeping the identity of treatment assignments masked for:

Subject	
Investigator (Treatment team/evaluator)	Single- blind ←
Monitoring committee (Sponsor)	Double- blind ←
	Triple- blind ←

Purpose of Blinding

Bias reduction
- Each group blinded eliminates a different source of bias.
- Blinding is most useful when there is a **subjective** component to treatment or evaluation.

Feasibility of Blinding

- **Ethics**: The double-blind procedure should not result in any harm or undue risk to a patient. (May be unethical to give "simulated" treatments to a control group, e.g., repeated injections.)

- **Practicality**: May be impossible to blind some treatments, e.g., radiation therapy equipment is usually in constant use; requiring a "sham" treatment might be a poor use of resources.

- **Avoidance of bias**: Blinded studies require extra effort (manufacturing look-alike pills, setting up coding systems, etc.). Consider the sources of bias to decide if the bias reduction is worth the extra effort.

- **Compromise**: Sometimes partial blinding (e.g., independent blinded evaluators) can be sufficient to reduce bias in treatment comparison.

Although blinded trials require extra effort, sometimes they are the only way to get an **objective** answer to a clinical question.

Reasons for Subject Blinding

If the treatment is **known** to the subject:
- Those on "no treatment" or standard treatment may be discouraged and drop out of the study.
- Those on the new drug may exhibit a placebo effect, i.e., the new drug may appear better when it is actually not.
- Subject reporting and cooperation may be biased depending on how the subject feels about the treatment.

Reasons for Treatment Team Blinding

Treatment can be biased by knowledge of the treatment, especially if the treatment team has preconceived ideas about either treatment, through:
- Dose modifications
- Intensity of patient examinations
- Need for additional treatment
- Influence on patient attitude through enthusiasm (or not) shown regarding the treatment

Reasons for Evaluator Blinding

If the endpoint is subjective, evaluator bias will lead to recording more favorable responses on the preferred treatment. Even supposedly "hard" endpoints (e.g., blood pressure, MI) often require clinical judgment.

Reasons for Monitoring Committee Blinding

Treatments can be objectively evaluated.

Recommendations to stop the trial for "ethical" reasons will not be based on personal biases.

(However, triple-blind studies are hard to justify for reasons of safety and ethics.)

Use of Placebos

The "placebo effect" is well documented.
> ⇒   When a new treatment is compared with **no** treatment, use of placebo is necessary to demonstrate a non-placebo effect.

Matched placebos are necessary so patients and investigators cannot decode the treatment assignment.

For example, in a trial of vitamin C for the common cold:
- Placebo was used, but was distinguishable upon breaking capsule open.
- Result: Many on placebo dropped out of the study. Those who **knew** they were on vitamin C reported fewer cold symptoms and duration than those on vitamin C who **did not know**.

Pill-matching is also useful in two-drug studies (drug A vs. B). Using a matched placebo for a specific drug, the following blinded treatment comparisons could be made:

Once-daily treatment		**vs.**	Twice-daily treatment		
Pill 1	=	slow-release tablet (200 mg)	Pill 1	=	conventional tablet (100 mg)
Pill 2	=	placebo	Pill 2	=	conventional tablet (100 mg)

Or, in a study of two completely different drugs, if the pills for Drug A and B cannot be made to look identical, then a placebo version of each can be used (each patient takes both pills, but one is a placebo):

Drug A		**vs**.	Drug B	
A =	active		A =	placebo
B =	placebo		B =	active

Bottle Coding

Pill bottles must be coded so that if one patient's assignment is discovered, it does not break the code for everyone.

For example, do **not** use:
- Even # = drug A
- Odd # = drug B

The code-breaking list must be readily available in case of a medical emergency.

Beware: People *will* try to break the code.

Assessing Whether Blinding Worked

At the end of the study, ask patients and investigators to guess which treatment the patient was on.

- If blinding worked, guesses should be 50% correct (with appropriate binomial variability).
- If the number of correct guesses > 50% $\Rightarrow$ can estimate how many knew: 2 x (proportion over 50%).*
- If the number of correct guesses < 50% $\Rightarrow$ suspect people know but are trying not to admit it by guessing the wrong treatment.

*To calculate the number of people who know their treatment assignment:

p = Proportion who know
$1 - p$ = Proportion who do not know

Expected proportion of correct responses: = $p + 1/2 (1-p)$

= $1/2 + (1/2)p$

For example, if we observe 75% correct responses:

0.75 = $1/2 + (1/2)p$
0.25 = $(1/2)p$
p = 0.50

$\Rightarrow$ 50% of people knew their treatment assignment.

Open-label Studies

Blinding (single or double), while desirable, may not be possible or feasible.

Open-label studies may be problematic, particularly in critical or pivotal studies. Bias may be reduced if there is a hard endpoint (such as survival or tumor size), but is particularly problematic if subjective scales are used.