CHAPTER VII

CLINICAL TRIALS

Francois Delahaye

Once the nature of the disease has been described and one is able to predict its natural course, the next question is: “what can be done for the patient?”

Faced with a patient presenting with a given illness, the doctor prescribes a treatment: medication, exercise, surgery, diet, etc.

But there are also many other ways of intervening to improve health: screening, prevention, education, at the level of the individual as well as at the level of the population.

Whatever the nature of the proposed intervention, the principle is to find out whether it does more good than harm to the patient or the community. The evaluation technique is the same regardless of the intervention: it is a clinical trial, which is a special case of a cohort study.

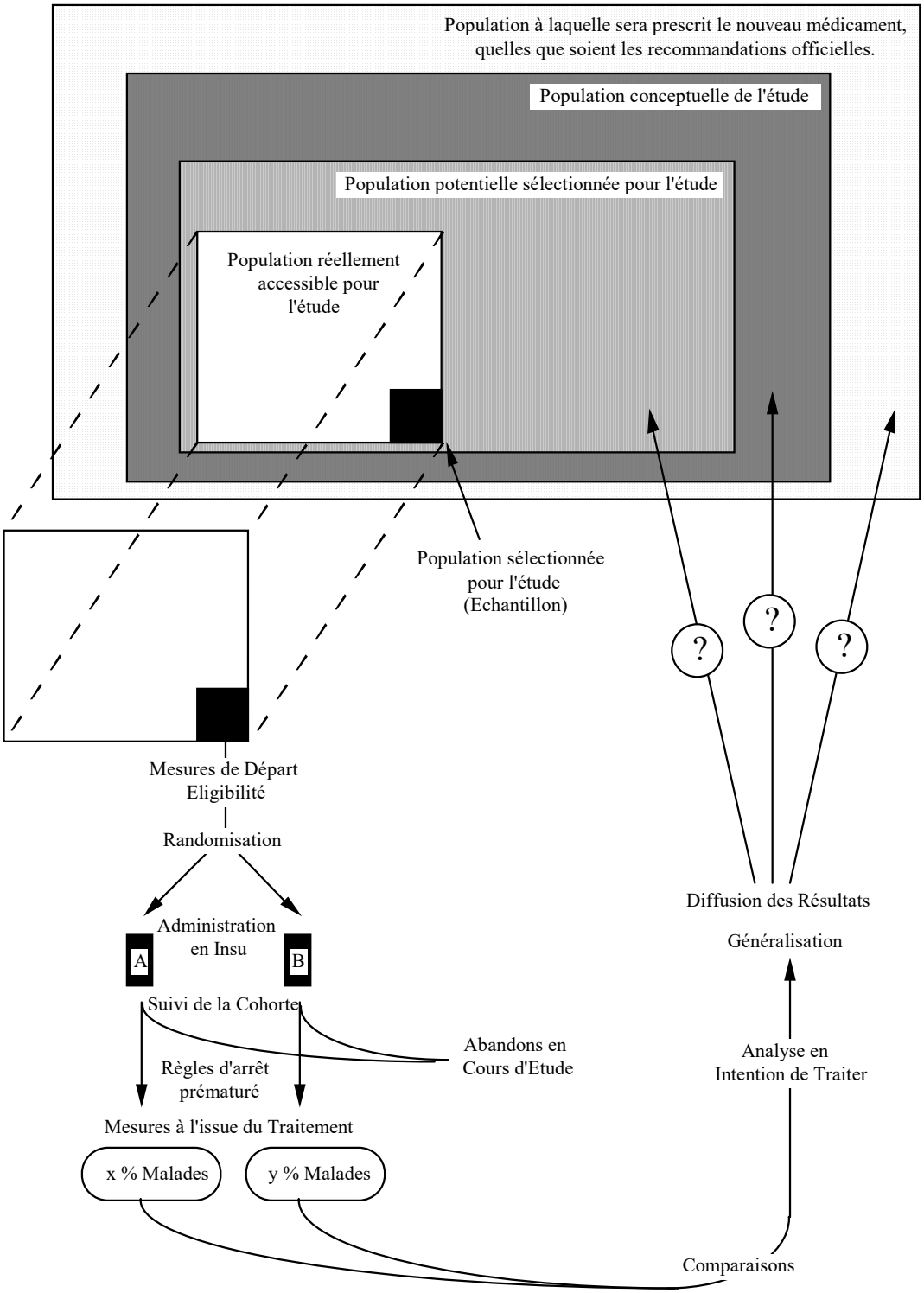
In a clinical trial, the conditions of the study—selection of groups to benefit from the intervention, nature of the intervention, management of follow-up—are under the control of the investigator. It is therefore an experimental study, similar in principle to those carried out in the laboratory. Its main advantage, compared to an observational study, is to present a better level of evidence when highlighting an association between two factors or a difference between two groups.

Clinical trials are cohort studies in which the investigator manipulates the factor being studied, such as a therapeutic intervention, and observes the effect on the endpoint. It belongs to one of the following five categories: death, disease, disability, discomfort and dissatisfaction (or their reverse).

The main advantage of experimentation over observational studies is the power of causal inference it allows (very high level of evidence).

Randomization—random assignment of subjects to the different groups—is the best way to control the influence of confounding factors, whether known or unknown, by distributing these factors equally among the different groups, so that their effect on the judgment criterion cancels out, and thus, if a difference is observed between the two groups, it is exclusively due to the effect of the factor studied (figure 1).

Figure 1 - General diagram of a therapeutic trial



Blinding avoids co-interventions.

The most frequent clinical trials are drug trials, which constitute a particular form of therapeutic trials, themselves classified in the category of intervention trials (alongside therapy, interventions can also be screening, prevention or of education; the test is carried out on the same model). In this chapter, we will mainly talk about therapeutic trials.

Therapeutic trials are essential for the development of new treatments. The first stage of development uses experimentation in vitro and on animals, and makes it possible to specify the pharmacology and toxicology of the product. Development then has four phases in humans:

- phase I studies aim to clarify safety and tolerance; they are made in a small number of subjects;
- phase II studies specify the optimal efficacy of the treatment;
- phase III studies establish the efficacy of the treatment, most often through comparative therapeutic trials, ideally randomised;
- Phase IV studies, after marketing, aim to establish any new indications and adverse effects not detected during the previous stages.

I - RANDOMIZED THERAPEUTIC TRIAL IN BLINDNESS

The development of a blind randomized therapeutic trial, or a 'randomized controlled trial' (RCT), comprises seven stages: the constitution of the study cohort, the realization of the basic measurements, the randomization, the administration of the therapy (the factor studied), patient monitoring, measurement and comparison of events in the different groups (the endpoint(s)), and finally the analysis of the results.

A - Constitution of the study cohort

This first step is to specify the characteristics of the population and the sampling method.

1 - Inclusion criteria

They define the main characteristics of potential and accessible populations (figure 1 above). The clinical and demographic characteristics (age, sex, race) make it possible to define the potential population, the one to which the results of the study can be generalized. The geographical and temporal criteria make it possible to define the accessible population, the part of the potential population which is available for the study. The definition of these criteria requires the acceptance of a compromise between scientific goals and practical constraints.

2 - Exclusion criteria

They are defined to eliminate, among the eligible subjects, those who risk interfering with the quality of the data or the interpretation of the results: alcoholics, patients with psychiatric problems, subjects likely to move house, etc. Exclusion criteria improve the feasibility of a study, but they must be used with moderation because a better homogeneity of the population studied is achieved to the detriment of the "generalizability" of the results. Some exclusion criteria are imposed by ethics, or by a subject's desire not to participate.

3 - Population

There are two main types of accessible populations:

- subjects from hospital-defined samples are inexpensive and easy to recruit, but selection factors can have a significant effect, especially on the generalizability of results to a large population;
- the subjects recruited at home constitute a representative sample of a specific region; these samples are particularly useful for guiding clinical or public health practice in the community; but there are two major disadvantages: the difficulty of realization and the high cost.

It is sometimes possible to avoid having recourse to sampling, and therefore to avoid biases, when the accessible population thus defined is truly accessible as a whole - this was the case in 1976 for the epidemic of legionellosis where all the cases been identified. This is the best approach. Usually, however, the accessible population is too large, so a smaller group must be selected: this is sampling.

4 - Sampling

Probability sampling uses random selection to ensure that each unit in the population has a specified probability of selection. This is the most rigorous approach.

It could be :

- a simple drawing of lots, for example using a table of random numbers generated by a computer (Table 1);
- a stratified draw (the population is divided into subgroups according to characteristics such as sex or age, and the draw is made in each of these groups);
- or a random draw of groups (natural groups of individuals are randomly drawn, for example football teams).

Non-probability sampling is much easier than probability sampling, but less rigorous, since it does not use chance.

Table 1 - Random number table

09801	20131	47650	20546	79800
01638	79004	13891	00746	26571
05441	02614	89720	18096	10974
58001	07467	19853	10074	32052
01985	49872	30106	24198	10023
14941	10123	45678	91019	51032
57489	32002	47921	00164	59758
74431	01320	48372	85967	45116
50206	12497	65773	12131	41516
17181	92021	22232	42526	27282
15424	70461	61241	21234	37989
15200	76746	59116	01246	42749
75975	46013	01654	97978	67240
10404	25704	01310	42795	79573
20275	12707	58067	84150	05178

B - Basic measurements

Some measures make it possible to characterize the subjects included in the study: surname and first name, address and hospitalization number, but also demographic and clinical characteristics, such as age, sex and diagnosis. These measurements are important informations because they also make it possible to compare the composition of the groups that make up the test. Usually, the first table in the final report of a RCT compares the various baseline characteristics of the subjects in the different groups. The goal is to verify that the differences are not greater than those that could have been observed by simple chance. If so, it would suggest a technical error in the randomization process. This would also run the risk, in the event of a difference observed between the two groups for the endpoint, that this difference is not due to the therapeutic intervention, but to the characteristic or characteristics for which an imbalance between the two is observed. groups.

It is often useful to measure the endpoint at the start of the study as well as at the end. When the criterion is a dichotomous variable (see Chapter 16), it is important to demonstrate that it is not present at the start. When dealing with a continuous variable, one can use the difference between the two groups in the degree to which the variable changed during the study. This approach controls for initial differences, and can give the study greater power than simply comparing values at the end of the study.

It is necessary to measure the various known predictors of the event studied (judgment criterion), that is to say the factors known to be able to have an influence on the judgment criterion, independently of the factor studied. This allows statistical adjustment of the results, which reduces the effects of an unexpected misdistribution of the predictor variables between the two groups. The efficiency of the study is thus increased. It also allows the investigator to examine these other predictors in another research question.

Finally, do not measure too many variables, as this increases cost and complexity.

C - Randomization

Random allocation establishes the basis for testing the statistical significance of differences between groups in the event of interest. It allows age, sex, and other baseline characteristics, known or unknown, that might “confound” an observed association to be distributed equally (barring random variation) among the randomized groups. Thus, the result observed between the groups at the end of the trial can be attributed to the effect of the intervention, since the effect that could have confounding factors — factors that modify the effects of the factor studied on the judgment criterion, because of their link both with the factor studied and with the judgment criterion — is equally distributed among the various groups.

The effects of a bad distribution by the simple fact of chance (on average 1 basic characteristic out of 20 is distributed differently between the groups at risk 0.05) are taken into account in the statistical tests making it possible to calculate the probability that chance is responsible for the difference observed between the groups for the studied event.

Because randomization is one of the cornerstones of a good RCT, it is important that it is well done. The two most important elements are that a true random allocation procedure be developed, and that the randomization process be unalterable, so that biases, intentional or not, cannot influence the process.

Usually, the patient undergoes basic investigations, is considered eligible for inclusion, and gives informed consent. He is then randomized, by the application, manual or automatic, of an algorithm to a set of random numbers, and his allocation in one of the groups is irreversible.

D - Administration of the various interventions

Randomization protects against the influence of confounding factors present at the time of allocation to the different groups. On the other hand, it has no effect on those that appear during monitoring.

Whenever possible, the investigator should design the mode of administration of the intervention in such a way that neither the subjects nor anyone in contact with them knows about the treatment received (Table 2). The term single blind is used when the patient does not know which product (tested intervention or placebo) he is taking; it is double-blind when neither the patient nor the doctor knows what the patient is taking.

Table 2 - The four levels of blinding

The blinding process can occur at four levels in a clinical trial:
1- The person in charge of allocating the patients to each group does not know how the patients already included in the study are allotted, so as not to risk changing the way they include the following patients in the study;
2- Patients do not know which treatment group they are in; thus, there is less risk of them changing their observance or describing their condition according to their membership in one or the other group;
3- The doctors in charge of monitoring the patients who participate in the study do not know to which group their patients belong; thus, their care is not likely to be modified, even unconsciously
4- When the researcher assesses the endpoint(s), he does not know to which group the patient belongs; thus, the measurement of the judgment criterion is not likely to be modified, even unconsciously.
The term "single-blind" refers to the patient alone, the term "double-blind" to the patient and the researcher

In an **open study** (without blinding), the investigator can pay particular attention to the patient when he knows that he is receiving the treatment being tested. This different attitude can represent a true intervention (co-intervention). Co-interventions can also affect the control group (e.g. subjects finding out they are on placebo request other treatments). These co-interventions may be the real reason for a difference in the frequency of the studied event between the groups. A partial solution to the problem of unplanned interventions is to specify and standardize the intervention.

A much more effective strategy is to make the study double-blind, that is to say to hide the nature of the treatment assigned, both to the subject and to the investigator. When double-blinding is technically sound, any unplanned intervention must affect both groups equally (with the exception, as with randomization, of chance misdistribution), and cannot alter the comparison of event between groups.

The logistical constraints can be heavy. You have to prepare identical capsules (shape, size, color, taste, etc.), and develop foolproof labeling and distribution systems.

It may be necessary to have a 24-hour unblinding mechanism in place when the situation demands that one be able to know very quickly which drug the subject is taking.

Another major difficulty is to ensure that neither the subjects nor the investigators can guess the assigned treatment.

Many therapeutic interventions cannot be performed without the knowledge of the doctor or the patient (e.g. surgery).

It is important to choose an intervention that can be generalized to daily medical practice. Choosing the right treatment can be particularly difficult in studies requiring several years of follow-up, because a treatment that was common at the start of the study may be outdated by the end of the study.

The best control groups are those who do not receive an active treatment, but a placebo identical in form, color, taste… to the drug studied. This strategy compensates for the possible placebo effect of the therapy tested, so that a difference between the study groups can be effectively attributed to a biological effect.

Another possibility is the comparison of a treatment with another treatment considered to be effective. If no difference is highlighted, the risk is to conclude that the two treatments are equivalent. In reality, the methodology of equivalence tests is different.

E - Patient Monitoring

A patient is observant (“compliant”) when he follows the instructions given by the medical profession on the treatment instituted or the intervention proposed. It is also referred to as treatment adherence. Patient compliance (for example, coming to scheduled consultations, and doing so on the scheduled date, taking the prescribed product, etc.) must be good.

The effect of the intervention, and therefore the power of the study, is reduced when the subjects are non-compliant. The investigator must try to choose an intervention that is easily tolerated, the single dose should be preferred.

Compliance must be measured, for example by self-administered questionnaire, tablet count or analysis of urinary metabolites. Non-compliance suggests a deliberate will on the part of the patient to disregard advice and prescriptions. But other factors can also come into play: the patient may misunderstand which medicine he should take, and at which doses, he may run out of medicine, not have money to go to the pharmacy or confuse the different medicines. …

F - Measurement of events or judgment criteria in the different groups and comparison

In choosing the type of endpoint, the investigator must often balance opposing considerations.

Often the events chosen as the endpoint of a study are not the true events, but surrogate events for the true phenomenon of interest, which limits the possible inferences (for example, in the trial of a fibrinolytic drug in the myocardial infarction, left ventricular ejection fraction or coronary patency rather than mortality can be used as a criterion for judging the efficacy of the product).

The measurement of the endpoint must be exact. Continuous variables have the advantage over dichotomous variables of increasing the power of the study, thus allowing the recruitment of a smaller number of patients. If a dichotomous variable cannot be avoided, the power depends more on the number of events than on the total number of subjects.

It is often desirable to have several variables measuring different aspects (for example, in the study of the efficacy of a drug in secondary prevention after myocardial infarction, the events studied may be mortality and recurrence of heart attack).

The investigator must also plan to measure the side effects related to the intervention, from the relatively minor symptom to the serious complication, even death. Assessing whether the benefits of an intervention outweigh its risks is the primary goal of most RCTs. Unfortunately, rare side effects are usually impossible to detect, regardless of the size of the study, and can only be discovered by a case-control study after the product has been widely used in the population.

The judgment criterion must be able to be measured without knowing the assignment of the subject to such and such a group.

The use of double-blinding is particularly important when measuring the endpoint requires subjective intervention on the part of the observer. The observer's knowledge of whether the patient belongs to the treated group or to the control group can modify his assessment of the endpoint. Double-blinding prevents the observation bias from being greater in one group than in the other.

The term triple blind is sometimes used to clearly show that the treatment taken is unknown to three people: the patient, the person administering the treatment, and the person measuring the endpoint.

The strategies for obtaining a high response rate are the same as those used in cohort studies: exclusion of subjects for whom monitoring seems difficult (alcoholics, psychiatric subjects, etc.); clear information to the subject on the importance of proper monitoring, elimination of those who find this monitoring difficult; recording the contact details of one or two relatives who will always know where the subject is, and those of the attending physician; regular telephone contact with patients.

Having monitoring for 100% or almost of the subjects can be essential when the event is rare and constitutes a possible cause of loss of sight.

G - Analysis of results

When the judgment criterion evaluated is dichotomous, the proportions of events in the groups are compared using the **chi2-test**.

When the event variable is continuous, a **t-test can** be used, or a **non-parametric test** when the variable is not normally distributed.

Time-aware methods are useful when there are differences in surveillance duration between participants, and **Cox-model regression analysis** can be used to adjust for uneven distributions of baseline confounders (which increases the power) (***Cf. Chapter XVI***).

Three issues should be considered when designing a study:

- **primacy of the analysis according to the intention to treat**, the only methodologically valid one: the results are analyzed, the patients remaining in the group to which they were assigned, even if they changed group during the trial (for example, a subject randomized to the surgical group but who is ultimately not operated on, or the reverse, a subject randomized to the medical treatment group but who is operated on afterwards);
- **ancillary role of subgroup analyzes** when they were not planned ***a priori*** in the protocol before the start of the trial: the subgroup studies carried out *a posteriori* are only there to generate hypotheses, which should be investigated in a new trial;
- **advisability of setting rules for premature termination** by providing for interim analyzes (with special statistical tests): it must be possible to terminate a trial prematurely if the intervention proves to be effective more quickly than expected (to allow all of the population concerned of this therapeutic advance) or on the contrary if it turns out to be harmful (so as not to expose the patients who receive it to it more).

II - SPECIFIC TYPES OF RANDOMIZED BLINDNESS TRIALS

A - Randomization after running-in period

It is useful for increasing the proportion of observant subjects. After identification of the study cohort and obtaining consent, all subjects are placed on a placebo. Later, those who have been observant are randomized. The exclusion of non-compliant subjects before randomization increases the power of the study and allows a better estimation of the effects of the intervention.

A variant of this plan study is the use, during the running-in period, of the active product. Here the response of an intermediate variable (i.e. which lies between the intervention and the event) can be used as a criterion for randomization (for example, for the study of the effect of an antiarrhythmic on mortality, we select patients in whom, during the running-in period, the antiarrhythmic has led to the disappearance of the arrhythmia).

It is important in the report of a run-in trial to specify any differences in baseline characteristics between randomized and non-randomized patients.

B - Factorial design

It allows multiple research questions to be answered in a single essay.

An example is the study of the influence of aspirin on myocardial infarction and that of beta-carotene on cancer (Figure 2). The subjects were randomized into 4 groups, and each of the two hypotheses could be tested by comparing two halves of the study cohort:

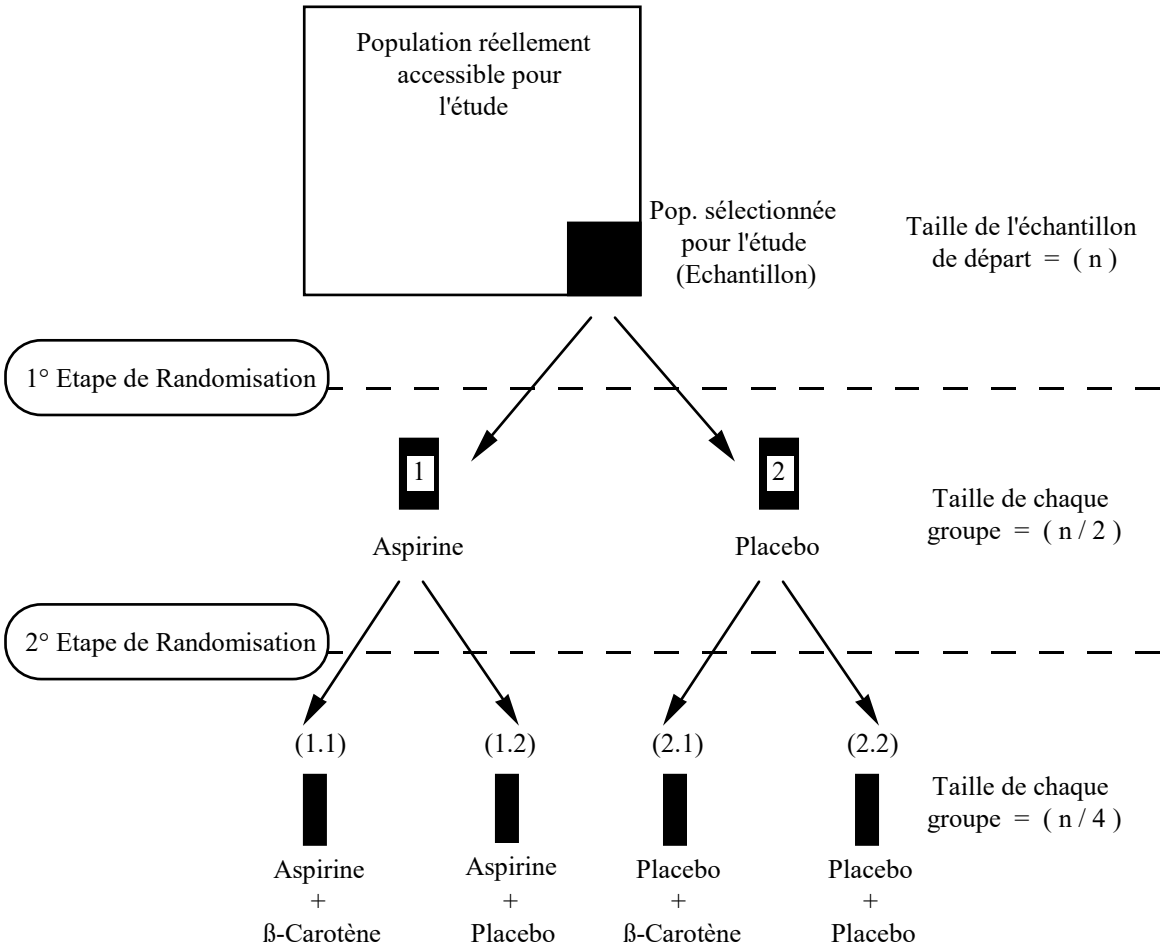
• all those taking aspirin are compared to all those taking placebo aspirin (regardless of the fact that half of each group receives beta-carotene);

• All those taking beta-carotene are compared to all those taking placebo beta-carotene (regardless of the fact that half of each group receives aspirin).

The investigator gets two complete studies for the price of one (no more subjects are needed than are needed for a trial testing only one hypothesis).

The factorial design is an extremely effective study design. The main limitation is the problem of the interactions between the two cause-effect relationships being studied.

Figure 2 - Blinded randomized therapeutic trial by factorial design Figure 3 - Study design in crossed permutations



C - Pairs of individuals

One strategy for distributing baseline confounders equally between the two groups is to select pairs of individuals, match them on important factors, such as sex and age, and then randomize each subject to one pair for each of the treatment groups.

A particularly appealing version of this approach is possible when the circumstances permit contrasting treatment and control effects in two parts of the same individual at the same time. For example, in subjects with diabetic retinopathy, each subject is randomly assigned one eye to treatment (photocoagulation), with the other eye serving as a control.

D - Pre-randomization

It assumes randomization before obtaining informed consent, which is then requested using different forms for the two groups.

This approach can increase the inclusion rate by removing the psychological barrier sometimes caused by the uncertainty of assignment to a treatment group, but the power is reduced in proportion to the proportion of subjects who refuse to be included, but which must be analyzed in order to satisfy the rule of analysis according to the intention to treat.

Furthermore, measuring the endpoint in subjects who refused to participate poses an ethical problem.

E - Randomization of groups

Instead of randomizing individuals, an investigator can choose to randomize natural groups of subjects, factories, cities...

Among the advantages of such a design, there is in particular the fact of avoiding that the subjects who receive a transmissible intervention, such as dietary advice, can discuss this intervention with acquaintances belonging to the same kind of population, but affected to the other group. But estimating study size and analysis is more difficult.

III - OTHER STUDY PLANS

A - Non-randomized clinical trials

They are much less satisfactory than the RCTs for controlling the influence of confounding factors. Analytical methods allow adjustment for unequally distributed background factors, but this does not solve the problem of unknown or unmeasured confounding factors.

B - Unblinded clinical trials

They are also less satisfactory than the RCTs, and prone to the risk of confusion due to co-interventions, and observational bias in events affecting one group more than the other.

When circumstances do not allow double blinding, single blinding is usually possible (the patient does not know what he is taking). However, this study design does not protect against co-interventions, and it should rarely be necessary: interventions that can be hidden from patients can usually also be hidden from investigators.

A more common form of partial blinding is the process of blinding the event (outcome) in an open-label study. Such studies can provide very useful conclusions, but these conclusions are usually less robust than those of double-blind studies.

C - Clinical trials in time series

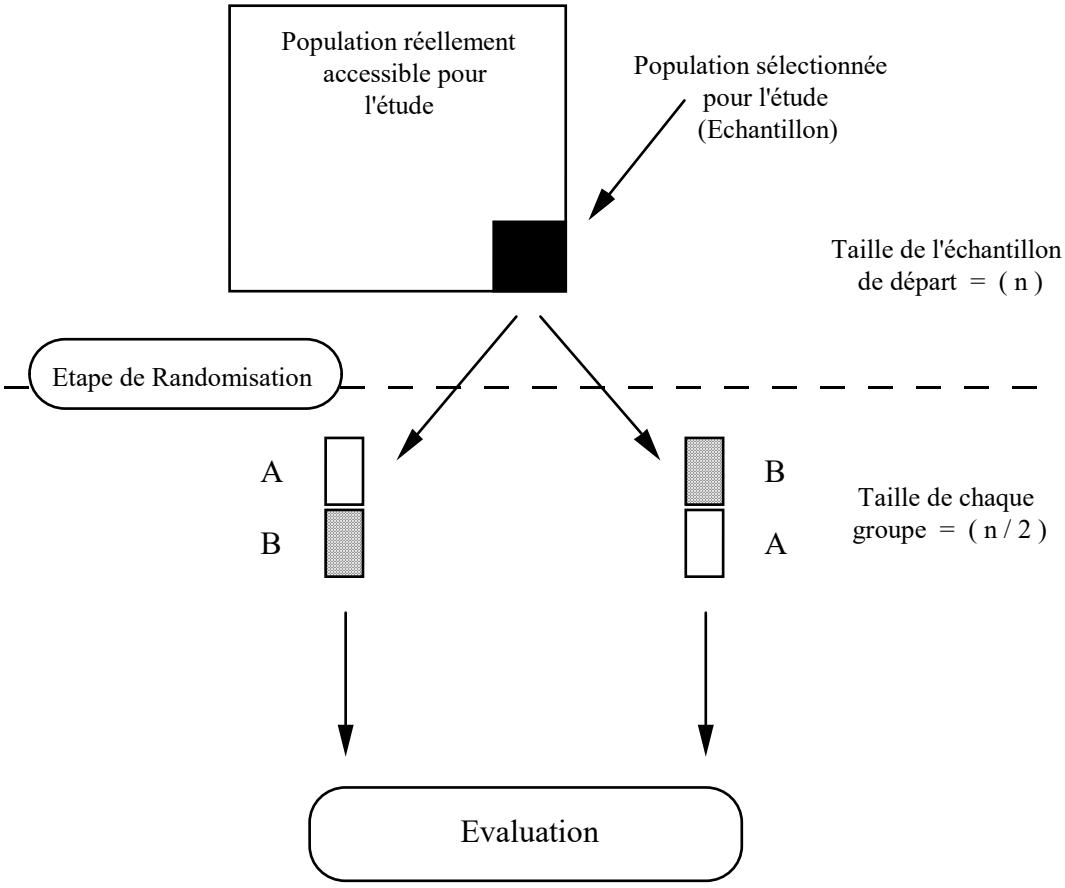
They can be useful for certain types of questions. Each subject is their own control during the sequential treatment and control periods. This means that personal characteristics such as age, sex, and genetic factors, potential confounders, are not evenly distributed, but simply eliminated. This also means that the study requires half as many subjects, since each subject provides both the control and experimental observations.

This study design is only useful in certain circumstances: studies in which the event responds rapidly and reversibly to the intervention or long-term experiments that cannot be randomized. The biggest drawbacks are the problem of confounding factors over time, and that of the carry-over effect (residual influence of the intervention on the event after this intervention has been stopped).

D - Study plan in crossed permutations

The influence of time-dependent covariates can be controlled by a crossover study design, in which half of the participants are randomized to receive placebo first, then treatment, and the other half to receive the reverse (figure 3).

Figure 3 - Study design in crossed permutations



This approach, or an equivalent such as the Latin square when there are more than two groups, has substantial advantages: control of confounding factors, effective doubling of the sample.

However, the disadvantages are often even greater: doubling the duration of the study, increased complexity of analysis and interpretation.

Cross-over studies are only a good choice when the subjects are difficult to recruit, when there is good reason to believe that the carryover effect is not a problem, or when this carryover effect constitutes by itself -even part of the research question.

E - Natural experimentation

The investigator analyzes a situation in which someone else applies an intervention. Natural experiments actually resemble observational studies as much as experiments since the intervention is not manipulated by the investigator, and the control for confounding factors is rather limited, unless the experiment includes a randomization.

IV - CONCLUSION

The double-blind randomized clinical trial is the gold standard in terms of clinical research protocol, providing the highest level of evidence when it comes to highlighting a relationship between two factors.

Randomization is the fundamental intervention since it eliminates the risk of error linked to confounding factors, and the blinding process eliminates the risk of bias linked to co-interventions.

The results of a trial are judged against two main questions:

- Is the treatment effective under ideal circumstances? The effectiveness of the treatment is judged in the patients who receive it, and who are fully cooperative, that is to say observant;
- is the treatment effective under ordinary conditions? The effectiveness of the treatment is judged in the patients who have been offered this treatment, and who are free to accept or refuse it; these are patients whose compliance may not be good; it is the difference between an experimental approach (ideal circumstances) and a pragmatic approach (ordinary circumstances or intention to treat); the latter approach offers results that are more easily generalizable.

It is obvious that it is not possible to answer all clinical research questions by a trial, for ethical, methodological or budgetary reasons. Analytical observational studies of the cohort or case-control type are then used. But their validity can be judged by the gap that exists between their protocol and that of a theoretical clinical trial that one would have liked to be able to carry out to answer the question.

In medical practice, a randomized clinical trial can disrupt doctor-patient relationships. The possibility for the patient to belong to the control group, the random allocation of treatment or

blind management can make the usual doctor-patient relationship, outside of experimentation, whose sole objective is the care of the patient, uncomfortable.

The clinical trial has been criticized for putting some patients in a situation in which they cannot benefit from the best possible treatment.

If indeed there is a good level of evidence to affirm the superiority of a treatment, not offering it to all patients is not ethical.

But if we really do not have proof of this superiority, then it is legitimate to offer the patient this treatment as well as its alternative.

One could even say that it is unethical to offer patients treatments that have not been rigorously evaluated and whose effectiveness is unknown.

On the other hand, although the randomized trial is expensive and difficult to conduct, the alternative to the trial, i.e. the administration of a treatment without solid information on its effectiveness, is probably much more expensive. Finally, a well-designed and well-conducted trial can save money.

It must therefore be considered that the principle of clinical experimentation is correct. It is only, but it is fundamental, to offer guarantees that a patient cannot participate in a trial against his will, especially in this doctor-patient relationship where the doctor can have considerable power. To propose to a patient to participate in a trial in such a way that he can refuse, and, if he accepts, with the guarantee that his rights will be respected, *these are* the legitimate constraints imposed by the ethics committees before which any protocol therapeutic trial must be submitted for approval. These committees ensure compliance with the fundamental principles of ethics in biomedical research: principle of the interest and benefit of research, principle of harmlessness of research, principle of respect for the individual, and principle of justice.

Patient information is fundamental. It is difficult to be sure of the fact that the patient signed the consent form really means that he understood all the information concerning the research. This means that the ability of the investigator to communicate fully and honestly with the patient is paramount.

References

Cucherat M. Methodology and interpretation of clinical trials. Editions Flammarion Medicine-Sciences, 2004

Friedman LM, Furberg CD, DeMets DL. Fundamentals of Clinical Trials. Editions Springer-Verlag, 1998