Phlébologie ANNALES VASCULAIRES 2020, 73, 3, p. 22-27

SOCIÉTÉ FRANÇAISE DE PHLÉBOLOGIE

Should randomized trials be the definitive clinical reference?

L'essai randomisé est-il la référence clinique absolue ?

Becker F. [1], Guex J.J. [2]

Article following an invited paper at the French Society of Phlebology Congress in December 2019.

Introduction

If by randomized trial one means the prospective, multicenter, double-blind, randomized controlled trial, then yes, the randomized trial tends to be the gold standard.

All authorities place meta-analyses of randomized controlled trials (RCT) at Level 1 evidence of efficacy, Grade A recommendation, i.e., as the optimal means of establishing evidence of superiority or non-inferiority of a treatment or strategy; they are the foundation, the backbone of recommendations and evidence-based medicine (EBM) [1].

However, each word has its importance, and a good understanding of these words, the care taken to avoid bias, the quality of the definitions, and the relevance of the evaluation criteria will determine the true value of a randomized trial.

We will successively consider the genesis of clinical trials, the different types of trials, the randomized controlled trial, the possible drifts and gimmicks, and the application to article reading.

Genesis of clinical trials, interest and limitations

The concept of a clinical trial probably dates back to Avicenna (980-1037) who already pointed out the need for a control group.

The history of controlled clinical trials begins in 1747 when Scottish naval surgeon James Lind (1716-1794) compared six remedies then used to treat scurvy, which killed many sailors on long voyages.

He demonstrated on 12 patients (2 patients per remedy) at similar stages of the condition, placed in the same part of the ship, receiving the same diet and care, that oranges and lemons, which we know today contain vitamin C, were the effective remedy for scurvy.

The Admiralty subsequently ordered that lemon juice be provided on all ships, which led to the disappearance of scurvy in the Royal Navy before the end of the 18th century[3].

Two centuries later, in 1948, the BMJ [2] published the first blinded randomized controlled trial demonstrating on 107 patients the benefit of Streptomycin in advanced pulmonary tuberculosis with a clear improvement in chest X-ray (51% vs *8%) and a significant reduction in mortality (7%* vs *27%).*

The rationale for the trial and the details of the procedure are remarkably well described, everything was said or almost said:

- problem well exposed (serious disease, product tested effective in vitro and on animals, some human tests not conclusive),
- low probability of spontaneous regression, need for a control group and ethics of control discussed (here streptomycin + bed rest 6 months *vs.* bed rest),
- randomization resulting in two groups differing only on the tested treatment, patients unaware of their participation in a therapeutic trial (informed consent is now required),

1. *Division of Angiology and Hemostasis, Geneva University Hospital, Geneva, Switzerland*
2. *Vascular Medicine, F-06510 Gattières.*
3. The story goes that this is one of the reasons for the defeat of the Franco-Spaniards in Trafalgar, despite their numerical superiority, the Franco-Spaniards had remained deficient in vitamin C and were less brave.

– utmost discretion about the trial (only the direct treating physician who received the envelope knowing the treatment).

A few years later (1955-1995), genius iconoclasts (T.C. Chalmers, A. Cochrane, D.L. Sackett, G. Guyat) were to move Medicine from empiricism, from dictates, from rote learning without discussion, to a practice based on solid validated data that would become Evidence-Based Medicine, EBM [3].

It was not without tensions and sometimes fierce criticism, but in 2020 EBM could be considered to have proved its worth and to have become a major key to medical reasoning. The media outbursts over hydroxychloroquine during the Covid-19 crisis have shown us that this is not yet the rule.

One could thus read "*The idea of returning to the '70s medicine, before the randomized double-blind placebo-controlled clinical trials is attractive",* and even *"one does not treat with statistics but with one's guts"* [4].

One can understand some criticism and a certain mistrust, or even rejection, of an overly dogmatic vision of EBM (it is not a question of replacing one dictatorship with another). There is no need for randomized trials for everything and not all situations have been evaluated by RCT, many commonplace situations are simply managed based on common sense and symptomatic treatment.

In emergency situations, especially when faced with the unexpected, one often manages with experience, pragmatism, and a few rules ... which does not prevent from validating this treatment, this attitude, afterwards as soon as possible.

For the rest, basing one's practice on well-conducted comparative trials is not at all incompatible with great medicine, quite the contrary.

– First of all, this does not exempt from listening and examining one's patient correctly and establishing the clinical probability a priori as low, intermediate or high [5].
– Then taking into account the placebo effect or the nocebo effect means taking into account the possible spontaneous favorable evolution of the clinical situation in question, the incidence of the doctor-patient relationship, the possible doubt/systematic rejection of the patient or the doctor.
– Concern for iatrogenic risk and the benefit-risk ratio means applying a key principle of medical ethics (*Primum non nocere*).
  Iatrogenic risk concerns not only therapeutics (whether drugs or invasive procedures) but also diagnosis (including cascading acts following non-justified procedures or the discovery of "incidentalomas").

It should also be taken into account:

– that a given drug may be effective in research laboratory and without clinical benefit,
– that the effect varies according to the dosage,
– that the improvement of an image or a physiological parameter does not ipso facto improve the patient,
– that a product may be effective in prevention and not in therapy, in secondary prevention but not in primary prevention,
– that what is effective in early stage of a disease is not necessarily effective in advanced stage (or vice versa), etc.

In order to assess the benefit-risk ratio, the cost-effectiveness ratio, to evaluate a new treatment, a new technique, a new strategy, it is necessary to have studies that address perfectly the given issue.

But first of all, a distinction must be made between retrospective and prospective studies, observational studies and randomized controlled studies.

Retrospective studies, whatever they are, even when paired, have intrinsic biases (we only deal with the content of the available records); they are useful for testing a hypothesis and often need to be verified by prospective studies.

Prospective studies are not free of bias, but they are the best way to proceed when evaluating a new treatment or strategy:

– either in a randomized cross-over trial (with the patient as its own control) if the study makes it possible,
– or in a randomized controlled trial against a reference treatment or technique or against placebo or abstention, making sur statistical power is optimal.

The purpose of randomization is to ensure that the groups are different only in the treatment or strategy tested; double-blind is intended to achieve maximum objectivity.

There are two types of RCT, explanatory trials and pragmatic trials.

– The first ones are conducted in expert centers and the question is "Is the intervention effective?".
– The second ones are conducted in any center treating patients responding to the given issue at hand and the question is "Is the intervention effective in everyday practice?".
  - *In explanatory trials* there is a strong pre-selection with numerous exclusion criteria in order to obtain a group with maximal expected effect and minimal adverse effects.
  - *In pragmatic trials*, on the contrary, pre-selection is weak as soon as the indication for the intervention is well respected.
  - *In explanatory trials*, the follow-up is strict and often rather heavy, the analysis can be done according to different modes, and the results are not always easily extrapolated in daily practice.

- *In pragmatic trials*, on the contrary, the comparator is often the usual management, the follow-up is close to the usual follow-up, only intention-to-treat analysis is performed (whether or not the patient has followed the treatment), and evaluation criteria are directly relevant for all the actors [8]. The advantage of a pragmatic RCT is that it is a study as close to real life as possible, with no or almost no exclusion criteria. The limitations are the difficulty to carry out the study blindly (especially double-blind) and to detect biases and confounding factors, and that it does not allow the effect of the different components of an intervention to be evaluated separately.

– The ideal is to start with an explanatory RCT and if positive to continue with a larger pragmatic RCT (Sackett).

So-called multicenter registry studies also have the advantage of studying the case in "real life" settings, but their value depends on the rigor brought to registry design and data collection.

It should also be noted that the results of well-conducted RCTs are only the truth of the moment until other trials show that a new treatment, a new strategy does even better.

Bias and drifts

Rules have been developed for the analysis and reporting of RCTs, particularly by the CONSORT group [6, 7].

The main purpose of these rules is to limit bias (an error that tends to produce a systematic difference between the measured effect and the actual effect, a distortion of the results systematically in favor or against the object of the study).

Among these biases are:

– *selection bias* which makes groups imperfectly comparable for general characteristics, risk factors and co-morbidities,
– *management and monitoring bias* (or performance bias)
– *and detection bias* (or criteria evaluation bias)

which must be prevented by standardization of practices and a perfectly respected double blind (when double-blinding is not possible, the evaluation must be carried out blind by a third party who is unaware of what has been done).

– *attrition bias* due to failure to account for all randomized and lost to follow-up patients, and protocol deviations or missing data [9].
– For those puzzled by the principle of the randomized controlled trial, let's highlight *a paradoxical bias* that must be addressed by a well-conducted randomization: participation in a trial can be an opportunity for the patient. In fact, in addition to the benefit of close monitoring, the mere fact of participating in the trial can improve it (Hawthorn effect, 10).

But as in many human activities, between the ideal, the theory and the practice, the reality on the ground... there are imperatives, slippages, manipulations, etc.

– One only needs to look at the small percentage of studies included in meta-analyses to be convinced of this.
– Moreover, despite the obligation to declare trials and the incentives to publish, it must be acknowledged that still nearly 50% of negative trials are not published [11].
– In addition, the impact of funding arrangements must also be questioned when industry-sponsored trials are more often favorable than independent trials [12].

It is therefore clear that there is great variability in the quality of these trials and the confidence that can be placed in them.

Another pitfall is the way we learn about clinical trials. If we do not perform a regular review of the literature, we most often become aware of clinical trials by

– the media (daily or weekly medical press or mainstream media)
– or through the follow-op of specialized journals.

One must be vigilant and have a critical eye.

– The presentation is too often too synthetic or in a too journalistic language with a quest for the "scoop" that can outweigh scientific objectivity.
– Every day we can see translation errors.
– Whatever the newspaper, one is attracted by the title of the article, and in specialized journals the lack of time leads too often the reader to go from the title to the conclusion of the abstract without reading the details [13].
– If the abstract, the totality of the abstract, holds our attention, we must go to the complete article starting with the chapter "material and methods" that we will read more or less in depth according to the interest.
– If the methods are correct we go to the results and the discussion before agreeing or not with the conclusion.

There are reading grids for the critical reading of articles [9, 14-16] which may seem complicated but are in fact quite simple once you get used reading an article diagonally in search of the various possible biases.

If the screening of the biases is negative, one takes the time to read the article in depth... there is only a small 10% left ...

Is the title of the article clear, and is the rationale and purpose of the study clearly stated in the introduction?

- Is this a prospective randomized controlled trial? (not a cohort or case-control study)?, is it an explanatory trial or a pragmatic trial?
- What about the control group: no treatment, placebo, reference treatment, other treatment? In case of a comparison of invasive treatments or techniques, is there

a medical group (optimal medical treatment, simple clinical management)?
- Is it double-blind (patient and doctor do not know which arm the patient is), single-blind (only patient does not know which arm he is) or open-ended (patient and doctor know in which arm is the patient)? A blinded study is not always possible; in this case, are the judgement criteria objective enough, has the evaluation been done by a third party who does not know the arm of randomization or even independent of the study?
- Is the case definition accurate and acceptable, consistent with clinical practice, neither too broad nor too narrow?
- Do the inclusion and non-inclusion (exclusion) criteria affect the representation of the population with respect to the issue being addressed? Is there any detail in these criteria that could lead to suspect a manipulation?
- Are the inclusion rate, number of patients included and duration of inclusion consistent with the prevalence of the addressed condition (beware of selection bias)?
- What about the ratio of included patients to patients likely to be included? Is it reasonable or does a rate that is too low suggest difficulties in inclusion?
- Are the primary (in principle there is only one) and secondary (there may be more than one) judgment criteria clinically relevant and clearly validated?
- Were the groups comparable at the beginning of the study, are they still comparable in the end?
- Were the two groups treated the same, except for the treatment being assessed?
- What is the rate of loss to follow-up? of discontinuing the study? of arm changes?
- Was the patient analysis an intent-to-treat analysis (analysis of all patients included in the trial according to the group in which they were randomized, regardless of the treatment actually received) or per-protocol analysis (analysis limited to the patients who actually received the treatment corresponding to the group in which they were randomized)? Intent-to-treat analysis is preferable to per-protocol analysis.
- Is the difference in results clinically relevant, both in absolute and relative terms?
- Is the final message consistent with the results achieved?
- Non-inferiority trials are becoming more and more frequent, what about the margin of non-inferiority, is it not excessive? If the result is expected one, does the discussion not evade the fact that it is a non-inferiority trial, perhaps suggesting then that it is a superiority trial?
- Also, do not forget to look at the conflicts of interest of authors and the funding of the study.

A few examples of drifts in vascular pathology:

- *Until recently, arterial pathology was attributed "on principle" to men and chronic venous pathology to women, which led to significant selection biases in the studies, even though epidemiological studies show that the prevalence of each of the arterial and venous affections is globally quite similar in both sexes ...*
- *Low-dose aspirin (75-100 mg/day) has long been used in cardiovascular prevention without much questioning*, until a distinction was made between primary and secondary prevention and randomised controlled trials showed that in primary prevention, low-dose aspirin is more harmful than useful, with an excess of major haemorrhages and no reduction in cardiovascular risk, except perhaps in very high-risk subjects for whom we can talk of primary-secondary prevention [17].
- *Carotid stenosis has given rise to a large number of publications and trials which are not free of bias and drift.*
  Firstly, there is still frequent reference to the ACAS study on the treatment of asymptomatic carotid stenosis (JAMA 1995), although it can be noted:
  - that it took 6 years to include 1,162 patients when 12,080 patients were operated on by the same people during this period and the patients were recruited mainly in vascular exploration laboratories and in vascular surgery (selection bias?),
  - that cardiovascular prevention (lifestyle, drugs) was far from the power of the current prevention mode,
  - that the study was prematurely stopped on a risk reduction (homolateral stroke, all strokes and death) of 50% but in fact an absolute risk from 2 to 1%,
  - that the benefit was for minor strokes and not major strokes.

  However, this presentation was followed by an explosion in the number of carotid endarterectomies for asymptomatic stenosis (10 times more). Since then, it has always been very difficult to have a proper medical arm in trials comparing endarterectomy and stenting, and medical treatment studies are struggling to find funding and inclusion.
  In endarterectomy vs. stenting studies the definition of asymptomatic carotid stenosis varies from study to study, studies of symptomatic stenosis end up including asymptomatic stenosis and we have even seen the asymptomatic increase in troponins included in a composite endpoint (ipsilateral stroke, all stroke-TIA, coronary events, elevated troponins, and death) to increase the rate of post-op cardiovascular events in the endarterectomy arm [18].
- *It makes sense to discuss venous compression in the prevention of post-thrombotic syndrome (PTS).*
  And the SOX study [19] has had a very strong impact since its publication in 2013 by showing the absence of difference between a placebo sock and a 30-40 mmHg sock in terms of the risk of PTS.
  A team experienced in clinical trials, a fine article in the Lancet.

However, some people immediately judged it harshly, even going so far as to ask whether the patients had worn the allocated sock [20].

Admittedly, compliance was mediocre to say the least (60% at 2 years), but in addition, this may be a good example of drift.

PTS was individualised for two sequelae which are difficult to treat, disabling venous claudication resistant to physical training (rather early and quite rare) and post-thrombotic ulcer (whose incidence is exponential up to 10-15 years after DVT).

Studies on the prevention of post-thrombotic ulcer have always come up against two major problems, the time of appearance of the ulcer (and therefore studies plagued by a high rate of loss of sight) and the very definition of post-thrombotic ulcer for the purposes of a study (how not to confuse it with an ulcer linked to a simple long saphenous reflux with incontinent distal perforator(s)?

To shorten the duration of the study the Villalta and Ginsberg scores were developed (in SOX the Ginsberg score was the primary endpoint and the Villalta score was a secondary endpoint), very good, but if these scores are sensitive they are not at all specific to PTS and the so-called venous symptoms, vesperic oedema, are common in the general population (it is therefore perhaps not surprising that there is no significant difference between the two groups on these criteria). It can also be noted that venous ulcer was not initially included in the Villalta score, that venous claudication was not considered, and that the rate of ulcers at 2 years in SOX equals the rate of ulcers at 10 years in the DURAC study [21].

More anecdotally, but perhaps not, it is said in the Introduction chapter of the Lancet article that compression socks are very difficult to put on and can cause discomfort (heat, constriction, skin irritation) and yet the majority of patients could not tell the difference between a placebo sock and a 30-40 mmHg sock, maybe the placebo sock has served its purpose, maybe it is the fact that the product was delivered by post, but this is surprising for anyone who has ever worn a 30-40 mmHg sock and/or it may prove those who say that the placebo compression sock does not exist ...

The end-of-year issue of the BMJ humorously illustrated this risk of drifting in the trials.

In 2003, a systematic review of the literature shows that there is no serious randomized controlled study demonstrating the effectiveness of the parachute, the effectiveness of which is in fact only based on observational studies (that's understandable...). In 2018, thunderbolt, a randomized controlled trial shows that wearing a parachute does not reduce the risk of death or serious trauma when jumping from an aircraft! Quid? in fact, if you look more closely you can see that the plane did not take off and that the study was carried out on the ground ... [22, 23].

To conclude:

Yes, the well-conducted prospective, multicentre, double-blind, randomised controlled trial, and even more so the meta-analysis of such trials, is the absolute reference, at least the reference of the moment until other study(s) shake(s) it/them up, nuance(s) it/them or do(es) it/them better.

But:

1. **The prestige of these essays does not dispense us from evaluating them with a critical eye before accepting their conclusions,**
2. **Even if they are flawless or almost flawless, they support the clinical judgement but do not replace it, they can be contested on a given case but then they must be justified.**

In fine, the well-conducted, prospective, multicentre, double-blind, randomised controlled trials forming the backbone of EBM, Sackett *et al.* clearly summarised the objective: "*Evidence-based medicine is the integration of best research evidence with clinical expertise and patient values*" [24].

Références

1. HAS. Niveau de preuve et gradations des recommandations de bonne pratique (2013).
2. Streptomycin treatment of pulmonary tuberculosis. A medical research council investigation. BMJ Oct. 30 1948.
3. Smith R., Rennie D. Evidence based medicine, an oral history. BMJ 2014 Jan 21;348:g371.
4. Claude Alexandre et Christian Pérronne in Haroche A. Raoult et la saga de l'hydroxychloroquine : les racines d'une passion. JIM Plus 30/05/2020.
5. Junod A.F. La décision médicale ou la quête de l'explicite. Editions Médecine & Hygiène 2007.
6. Schultz KF, Altman DG, Moher D. for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c332.
7. Boutron I., Moher D., Altman DG., Schultz KF., Ravaud P. for the CONSORT Group. Methods and Processes of the CONSORT Group: example of an extension for trials assessing nonpharmacologic treatments. Ann Intern Med. 2008 ;148 :W-60-W-66.
8. Michiels B., Eerstelinjs V. Quelle est la grande particularité des essais cliniques pragmatiques ? Minerva déc. 2014.
9. Mismetti P., Laporte S. Méthodologie des essais cliniques : pourquoi et pour qui ? Les notions essentielles à une lecture critique. Médecine thérapeutique, Vol 13, N° 4, juillet-août 2007.

10. McCarney R, Warner J., Iliffe S., van Haselen R., Griffin M., Fisher P. The Hawthorne effect: a randomised, controlled trial. BMC Med Res Methodol 2007;7 :30.
11. Goldacre B., DeVito NJ., Heneghan C., Irving F., Bacon S. et al. Compliance with requirement to report results on the EU Clinical Trials Register: cohort study and web resource. BMJ 2018;362:k3218.
12. Lundh A., Lexchin J., Mintzes B., Schroll JB., Bero L. Industry sponsorship and research outcome (Review). Cochrane Library, Feb 2017.
13. Mandell BF. Clinical trials : More to learn than the results. Clev Clin J Med 2019 Jul, Vol 86, Nr 7.
14. Perrier A. L'essai randomisé contrôlé, un guide de lecture. Rev Mal Respir 2002, 19, 505-14.
15. Pariente A. Analyser et utiliser les résultats des études cliniques. Rev Prat 2013 nov, 63, 1305-10.
16. Cucherat M. La lecture critique des essais cliniques http://www.spc.univ-lyon1.fr/lecture-critique
17. Ravach G., Fournier S., Mazzolai L., Nanchen D. Aspirine en prévention cardiovasculaire primaire : la fin d'une époque ? Rev Med Suisse 2020; 16: 459-62.
18. Paraskevas KI., de Borst G., Veith FJ. Why randomized controlled trials do not always reflect reality. J Vasc Surg 2019;70:607-14.
19. Kahn SR., Shapiro S., Wells PS., Rodger MA., Kovacs MJ. et al for the SOX trial investigators. Compression stockings to prevent post-thrombotic syndrome: a randomised placebo-controlled trial. Lancet published online december 6, 2013.
20. Laproboulos N., Gasparis A., Caprini JA., Partsch H. Lancet 2014, 34 July 12. Correspondance.
21. Schulman S., Lindmarker P., Holmstro M., Carlsson A., Nicol P. et al. Post-thrombotic syndrome, recurrence, and death 10 years after the first episode of venous thromboembolism treated with warfarin for 6 weeks or 6 months. J Thromb Haemost 2006; 4:734-42.
22. Smith GCS., Pell JP. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. BMJ 2003;327:1459-61.
23. Yeh RW., Valsdottir LR., Yeh MW., Shen C., Kramer DB. et al on behalf of the PARACHUTE Investigators. Parachute use to prevent death and major trauma when jumping from aircraft: randomized controlled trial. BMJ 2018;363: k5094.
24. Sackett DL., Strauss SE., Richardson WS., Rosenberg W., Haynes RB. Evidence-based medicine: how to practice and teach EBM. 2nd ed. Edinburgh, Scotland: Churchill Livingston; 2000.