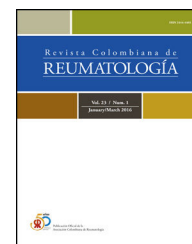




Revista Colombiana de REUMATOLOGÍA

www.elsevier.es/rcreuma



Review article

Clinical trials: General concepts and interpretation



Pablo Castaño González^a, José A. Gómez-Puerta^{b,*}

^a Artmédica, Medellín, Antioquia, Colombia

^b Servicio de Reumatología, Hospital Clínic, Barcelona, Spain

ARTICLE INFO

Article history:

Received 16 November 2018

Accepted 9 November 2021

Available online 7 July 2023

Keywords:

Clinical trial

Randomisation

Controlled studies

Palabras clave:

Ensayo clínico

Aleatorización

Estudios controlados

ABSTRACT

There are different types of clinical trial designs that serve specific purposes, which will depend on the nature of the rheumatological disease, the characteristics of the therapeutic interventions being investigated and the question the study is intended to answer. This review describes the main types of clinical trials conducted in the field of rheumatology, including their applications and limitations.

© 2021 Asociación Colombiana de Reumatología. Published by Elsevier España, S.L.U. All rights reserved.

Ensayos clínicos: conceptos generales e interpretación

RESUMEN

Existen diferentes tipos de diseños de ensayos clínicos que sirven a propósitos específicos y dependerán tanto de la naturaleza de la enfermedad reumatológica como de las características de las intervenciones terapéuticas investigadas y la pregunta que se desea responder con el estudio. En la presente revisión se describen los principales tipos de ensayos clínicos que se realizan en el área de la reumatología, incluyendo sus aplicaciones y limitaciones.

© 2021 Asociación Colombiana de Reumatología. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Introduction

With the appearance of biological therapies two decades ago, the presence of clinical trials that assess the efficacy or the safety of these and other therapeutic strategies has increased in the literature of the specialty; therefore, it is important for

rheumatologists to know the basic concepts on the types of clinical trials at the time of interpreting them, designing them or participating in their implementation.

The different modalities of clinical trials include: single-arm trials, placebo-controlled trials, crossover trials, factorial trials, non-inferiority trials, and studies designed to evaluate new diagnostic devices, among others.

DOI of original article: <https://doi.org/10.1016/j.rcreu.2021.11.004>.

* Corresponding author.

E-mail address: jagomez@clinic.cat (J.A. Gómez-Puerta).

2444-4405/© 2021 Asociación Colombiana de Reumatología. Published by Elsevier España, S.L.U. All rights reserved.

The selection of the type of study depends on the clinical question that is intended to be answered, the characteristics of the disease and the therapeutic interventions that are evaluated, the outcomes selected, the availability of a control group and the time and economic resources that are available for its development. The different phases and characteristics of the main clinical trial designs used in rheumatology are described below.

Phases of clinical trials

When clinical trials evaluate the effect of a new therapeutic intervention that aims to receive approval from the drugs and medical devices regulatory agencies, a classification by phases of research in humans is used. Phase I studies usually involve less than 100 participants (generally healthy volunteers) and their objective is to investigate safety, tolerability, pharmacokinetics and pharmacodynamics. Phase II studies are generally conducted with less than 300 individuals with a specific disease and seek to evaluate the biological or clinical effect of the safe dose found in phase I and to continue investigating its safety profile. Phase III studies are usually multicenter randomized clinical trials, with a large number of patients with a determined pathology (between 300 and thousands), whose objective is to confirm the efficacy and safety in relation to the standard therapy for that condition (or the placebo, if such therapy does not exist) and thus get marketing approval. Phase IV or post-marketing surveillance studies aim to continue evaluating the safety of the drug or device over a longer period and in a number of individuals larger than in phase III studies.¹

Single-arm trials

In this type of trial, a sample of individuals with a specific disease is exposed to a determined therapy and followed-up during a period to assess the response thereof, that is, all participants receive the experimental treatment.

This kind of study allows to obtain preliminary data on the efficacy of an experimental intervention and, despite it is the simplest to design, the interpretation of the results should be cautious, because is difficult to know if there is a *placebo effect of the therapy used*, if the response obtained is due to the changes inherent to the natural history of the disease, as well as to establish the magnitude of the response without a control group to compare. Due to these limitations, the single-arm trials should be conducted in well-selected diseases, in which the placebo effect is null or minimal, the natural history is known and it is not possible, difficult or unethical to use a control arm. Its main application is in rare or orphan diseases, in which the number of patients is scarce; on the other hand, these trials are not useful in patient with osteoarthritis, fibromyalgia and other conditions of chronic pain in which there are high rates of placebo effect.²

For example, a single-arm clinical trial investigated the safety and efficacy of canakinumab (anti-IL-1 β human monoclonal antibody) in Schnitzler syndrome, a rare autoinflammatory disease characterized by chronic urticaria, monoclonal gammopathy, and clinical manifestations of systemic inflam-

mation such as intermittent fever, arthralgia or arthritis, bone pain, adenopathies and hepatosplenomegaly, and in which there is a higher risk of developing lymphomas and AA amyloidosis. In this trial were included 8 patients who received canakinumab at a dose of 150 mg subcutaneously every month, for 6 doses, and were followed-up for 9 months. The primary outcome was complete remission on day 14, which was achieved by all patients; 3 months after the therapy with canakinumab was discontinued, 4 patients relapsed, while 2 persisted in remission 7 months after the last dose.³

Placebo controlled trials

In this type of study, a sample of individuals with a specific disease is selected and randomized into 2 or more groups: one or more that receive an experimental treatment (it can be either the same at different doses or different treatments) and another that receives placebo. The individual randomized receives only one type of treatment throughout the duration of the trial. Subsequently, the individuals are followed-up during the study and the results are compared.⁴

Many clinical trials in rheumatology use this design. An example is the phase II randomized, placebo-controlled, multicenter, double-blind clinical trial, in which the efficacy and safety of apremilast (oral phosphodiesterase 4 inhibitor) were investigated in 111 patients with Behçet's disease with 2 or more active oral ulcers, who were randomized into 2 groups: apremilast 30 mg every 12 hours and oral placebo every 12 h; the duration was 12 weeks, and after this period there was an extension phase of active treatment (phase in which the patients who received placebo were switched to receive treatment with apremilast) for 12 weeks. The primary outcome was the number of ulcers after the first 12 weeks, and the results favored apremilast, which motivated the implementation of a phase III clinical trial: the Relief study, which included a greater number of patients (207) and a longer extension phase of active treatment (52 weeks), in which the benefits of this drug in the treatment of oral and genital ulcers were confirmed.

Placebo-controlled trials in which the various interventions have been randomly assigned allow to control for both known and unknown confounding factors, after an intention-to-treat (ITT) analysis.

In the ITT analysis, patients are analyzed according to the intervention to which they have been initially assigned (randomized patient, analyzed patient, regardless of whether or not they received the intervention). In the case that the samples are well balanced after randomization (they tend to be so in most clinical trials), the comparisons made between both interventions (experimental therapy vs. placebo) will more accurately reflect the magnitude of the effect of the medication or intervention on the disease selected for the clinical trial.⁵ It should be taken into account that stratified randomizations are made in some studies, generally based on knowledge from previous studies, to balance the samples or to select the group to be studied (for example, including 3 patients with seropositive rheumatoid arthritis for each patient with seronegative rheumatoid arthritis, one male patient with systemic lupus erythematosus [SLE] for every 9 women, etc.).⁶

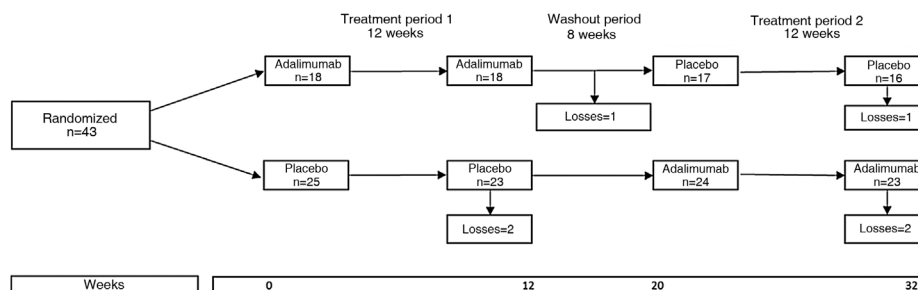


Fig. 1 – The HUMOR study illustrates the characteristics of a crossover clinical trial: random assignment to ADA and placebo in period 1, a washout period, and the crossover of the groups in period 2.

Source: modified from Aitken et al.⁸

On the other hand, in the per-protocol analysis, only those patients who received at least one dose of the intervention or treatment are included and are usually used to assess the safety of a drug or device; for this reason these analyses may be affected by the loss of patients during follow-up.

The disadvantage of these parallel trials (trials in which each patient receives throughout the study the intervention to which was initially assigned) is that they require a large sample size, due to the variability in the response of each patient and each treatment; if it is required to detect small differences, the sample must be larger.

Trials with crossed groups (crossover)

In this type of design each participant is randomized to a determined sequence of treatments that will be administered during specific periods of treatment, separated by a period of washout of the effect of the predecessor intervention, that is, each individual acts as his/her own control because he/she consecutively receives each of the study treatments at different times. For example, in a 2×2 crossover trial each participant is randomized into one of the two sequences of treatments: A and then B, or B and then A. The washout period, also known as stabilization period, aims to eliminate the probable residual effect after the suspension of the first treatment administered and, in case that the compared strategies are drugs, it is recommended that it be at least 5 times the half-life of the study drug with the longest half-life (It is not useful in therapies with long residual therapeutic effects such as rituximab or zoledronic acid).

Since each subject serves as his/her own control, these studies have different advantages compared to parallel designs (those in which each participant receives only one of the interventions under study): reduced sample size, less interindividual variability, and a greater probability of recruiting patients, since each participant will receive at least one active treatment at some point. Their main limitations are: a higher risk of loss to follow-up due to the exposure of each individual to 2 or more therapies, which implies a longer duration of the study and a higher probability of side effects and death; the residual effect: persistence in a period of the effect of the intervention of the previous period; the sequence effect: the order of the interventions affects the final result and only the effect of the first one can be reliably assessed; the period

effect: the basal characteristics of a patient change during the study, so that during the second or subsequent periods they are not similar to those at the beginning; the impossibility of evaluating long-term efficacy and safety. For these reasons, this design is useful when investigating therapies with a rapidly reversible effect after discontinuation or a short half-life, with a rapid onset of effect that allows short treatment periods and in chronic diseases with relatively stable clinical manifestations over time and low risk of losses (for example, it would not be very useful in SLE).⁷

An example of clinical trial with crossover groups is the HUMOR (*HUMira for erosive hand Osteoarthritis*) study, in which the efficacy of adalimumab (ADA) was evaluated in 51 patients over 50 years of age who met the criteria of the American College of Rheumatology for osteoarthritis of the hands and who had at least one erosion by conventional radiograph and synovitis by magnetic resonance imaging.⁸ Each participant was randomized to receive ADA 40 mg subcutaneously every 2 weeks or subcutaneous placebo for 12 weeks, followed by a washout period of 8 weeks and then was crossed to the other treatment for another 12 weeks. Expressed in another way, each patient was randomized into two groups or sequences: group 1, which received the active treatment and then placebo, and group 2, which received placebo and then the active treatment (Fig. 1). The primary outcome was the change in the visual analog pain scale at 12 weeks; the results were negative because there was no significant difference between the groups.

Factorial clinical trials

In this type of clinical trial, 2 or more experimental interventions are evaluated simultaneously, each one separately and in combination. In the simplest design (2×2) there would be 2 treatments or factors: A and B, and the arms of the clinical trial would be 4: A alone, B alone, A + B and none of them. The greatest efficiency of these studies occurs when it is assumed that the treatments under study have no significant interaction, that is, the effect of treatment A does not depend on whether treatment B is administered, and vice versa: nevertheless, these studies can also be used when it is wanted to study specifically if there is an interaction between the treatments.

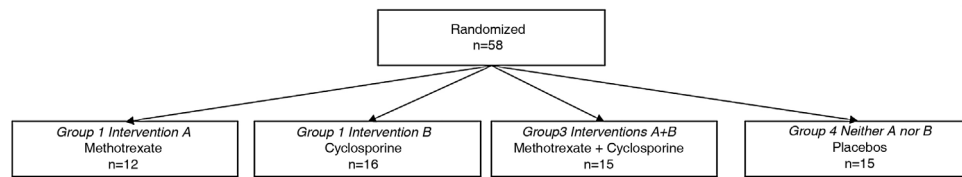


Fig. 2 – The SELAM study is an example of the elements of a factorial clinical trial. In group 1 of the figure, both interventions (MTX and cyclosporine) were used, in groups 2 and 3 only one of the interventions (MTX and cyclosporine, respectively) and in group 4 none of these (oral placebos).

Source: taken from Ibrahim et al.⁹

The SELAM study⁹ was a placebo-controlled factorial clinical trial investigating the efficacy of two treatments: cyclosporine (A) and methotrexate (MTX) (B), during 56 weeks, in 58 adults with dermatomyositis or polymyositis with incomplete response to glucocorticoids. There were 4 arms in the study: MTX 15–25 mg/week+glucocorticoids (A), cyclosporine 1–5 mg/kg/day+glucocorticoids (B), MTX+cyclosporine+glucocorticoids (A+B) and placebo of MTX+placebo of cyclosporine+glucocorticoids (neither A nor B); note that steroids in this case are a co-intervention present in all groups (Fig. 2). The primary outcome was the score on the manual muscle test at 12 months. In all groups there was an improvement in outcome, but no significant differences were found when comparing each of the active treatment arms (alone or in combination) with the placebo arm.

Non-inferiority trials

Non-inferiority trials are used when it is desired to study whether the effect of an experimental intervention is non-inferior compared to that of a standard intervention that has previously demonstrated to be superior to placebo. The rationale for using this design lies in the fact that the experimental

treatment that seeks to demonstrate to be not inferior has a potential advantage over the standard treatment in terms of safety, cost, administration, dosage, tolerability, or some other characteristic that is convenient or advantageous.¹⁰ To exemplify this, we can mention the case of biosimilar medicines, which are compared with the innovative biologic to demonstrate non-inferiority, but not superiority, and whose main advantage should be a lower cost.

The fundamental statistical strategy in this type of study is the non-inferiority margin or non-inferiority criterion, which is derived from the results obtained in a clinical trial of superiority of the standard treatment previously performed or from a meta-analysis of several of these. The results of a non-inferiority clinical trial can be: 1) demonstrated non-inferiority: the experimental treatment is non-inferior to the comparator; 2) non-demonstrated non-inferiority: the experimental treatment cannot be considered non-inferior to the comparator (in other words, it is inferior); 3) superiority: the experimental treatment is superior to the comparator, and 4) non-conclusive: the effect of the comparator treatment exceeds the margins of inferiority predetermined by the study, but the confidence interval includes values within the desired effect of the drug in relation to the comparator; this usually happens when the sample size was not sufficient or a

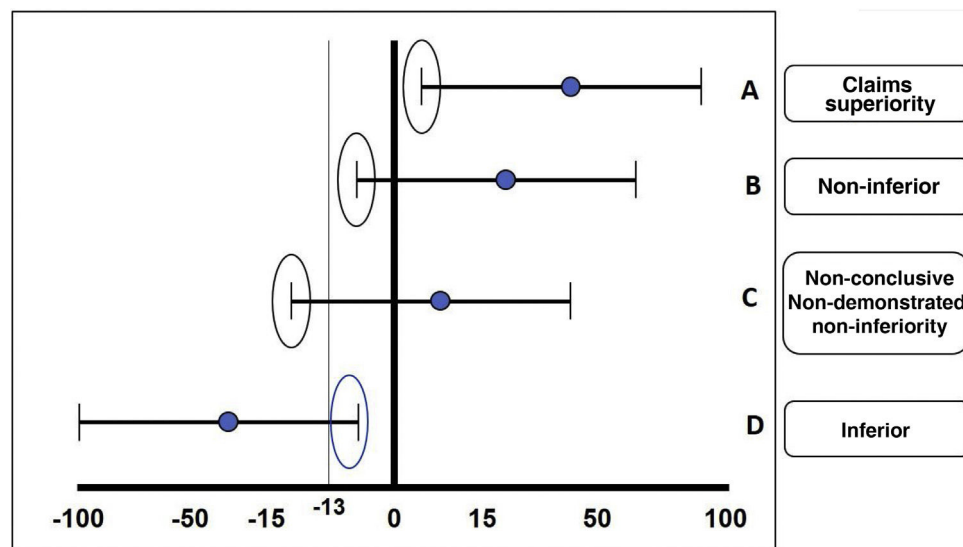


Fig. 3 – The different possible results in the case of an inferiority trial are illustrated: A) Claims superiority. B) Non-inferiority. C) Non-conclusive. D) Inferiority. In this case, a lower margin of -13 is taken as an example.

Source: based on the design of the ORAL Strategy study.¹¹

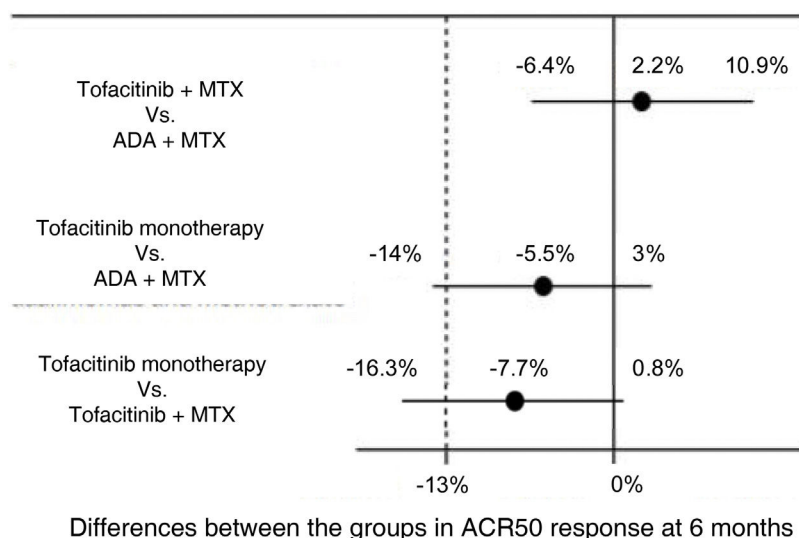


Fig. 4 – Results of the ORAL Strategy according to the three arms of the study.

Source: modified from Fleischmann et al.¹¹

very ambitious non-inferiority margin was calculated for that molecule or intervention (Fig. 3).

For example, the ORAL Strategy study¹¹ was a randomized, controlled, double-blind, phase 3b/4, head to head clinical trial with a non-inferiority design in which the efficacy of three therapeutic strategies in patients with active rheumatoid arthritis with inadequate response to MTX was evaluated: tofacitinib in monotherapy, tofacitinib plus MTX and ADA plus MTX. The experimental interventions are the groups of tofacitinib, while the active comparator standard treatment is the group of ADA plus MTX. The primary outcome was the proportion of patients who achieved at least an ACR 50 response at 6 months. The non-inferiority margin was -13%, percentage that represented half of the absolute difference observed between ADA plus MTX in relation to placebo—which was obtained from a metaanalysis of clinical trials on ADA,¹²—and the midline was 0. Non-inferiority was declared if the lower limit of the confidence interval for the difference between the groups was greater than -13% and superiority if the lower limit was greater than 0. The results of the study are illustrated in Fig. 4.

Another non-inferiority design is the study of lupus nephritis, which was an open-label randomized clinical trial comparing intravenous cyclophosphamide (CFM) versus mycophenolate mofetil (MMF) during 24 weeks as a remission induction treatment in 140 adults with proliferative lupus nephritis (classes II and IV).¹³ The main argued motivation to conduct the study was the assumption that MMF would have less adverse effects and better tolerance than CFM, the standard therapy at that time. The primary outcome was the complete remission rate at 24 weeks. The non-inferiority margin was set by the researchers at -10% (a single tail). The criterion to establish non-inferiority was that the lower limit of the confidence interval of the absolute difference in remission rates between the MMF group and the CFM group was greater than -10%, while the criterion for superiority was that it would be greater than 0. This difference was 16.7%, with a 95% con-

fidence interval (95% CI) of 5.6%–27.9%; the lower limit of this interval not only was greater than -10%, but also crossed 0, for which superiority of MMF was assumed, which constitutes an example of a non-inferiority trial in which superiority is claimed.

Conclusions

There are different designs of clinical trials that serve specific purposes, which depend on the nature of the rheumatic disease, the characteristics of the therapeutic interventions investigated, and the question to be answered with the study. Given the large number of publications of clinical trials, it is the responsibility of the clinician to know how to interpret these results and understand the methodology of the study, as well as its validity, in order to finally give it the applicability that he deems appropriate with a specific patient.

Conflict of interest

Dr. Gómez Puerta has presented lectures for laboratories that conduct clinical trials, including Abbvie, Bristol-Myers Squibb, Galápagos, Janssen, Lilly, MSD, Pfizer and Roche. None of the mentioned laboratories made a contribution for the preparation of this manuscript.

REFERENCES

1. Friedman LM, Furberg CD, DeMets DL. Introduction to clinical trials. In: Fundamentals of clinical trials. New York, NY: Springer New York; 2010. p. 1–18.
2. Evans SR. Clinical trial structures. J Exp Stroke Transl Med. 2010;3:8–18, <http://dx.doi.org/10.6030/1939-067x-3.1.8>.
3. De Koning HD, Schalkwijk J, van der Ven-Jongekrijg J, Stoffels M, van der Meer JWM, Simon A. Sustained efficacy of the monoclonal anti-interleukin-1 beta antibody canakinumab in

- a 9-month trial in Schnitzler's syndrome. *Ann Rheum Dis*. 2013;72:1634-8, <http://dx.doi.org/10.1136/annrheumdis-2012-202192>.
4. Katz DL. Study design. In: *Clinical epidemiology & evidence-based medicine. Fundamental principles of clinical reasoning & research*. Sage Publications; 2001. p. 147-70.
 5. Capurro ND, Gabrielli NL, Letelier SLM. Importancia de la intención de tratar y el seguimiento en la validez interna de un estudio clínico randomizado. *Rev Med Chil*. 2004;132:1557-60.
 6. Molina Arias M, Ochoa Sangrador C. Ensayo clínico (III). Aleatorización. Enmascaramiento. *Evid Pediatr*. 2015;11:15.
 7. Rothman KJ. Epidemiology in clinical settings. In: *Epidemiology: an introduction*. 2nd edition Oxford University Press; 2012. p. 235-53.
 8. Aitken D, Laslett LL, Pan F, Haugen IK, Otahal P, Bellamy N, et al. A randomised double-blind placebo-controlled crossover trial of HUMira (adalimumab) for erosive hand Osteoarthritis – the HUMOR trial. *Osteoarthritis Cartil*. 2018;26:880-7, <http://dx.doi.org/10.1016/j.joca.2018.02.899>.
 9. Ibrahim F, Choy E, Gordon P, Doré CJ, Hakim A, Kitas G, et al. Second-line agents in myositis: 1-year factorial trial of additional immunosuppression in patients who have partially responded to steroids. *Rheumatology (Oxford)*. 2015;54:1050-5, <http://dx.doi.org/10.1093/rheumatology/keu442>.
 10. Ferreira-González I. Bases para la interpretación de los estudios de no inferioridad: a propósito de los estudios ROCKET-AF, RE-LY y ARISTOTLE. *Rev Esp Cardiol*. 2014;67:432-5.
 11. Fleischmann R, Mysler E, Hall S, Kivitz AJ, Moots RJ, Luo Z, et al. Efficacy and safety of tofacitinib monotherapy, tofacitinib with methotrexate, and adalimumab with methotrexate in patients with rheumatoid arthritis (ORAL Strategy): a phase 3b/4, double-blind, head-to-head, randomised controlled trial. *Lancet*. 2017;390:457-68, [http://dx.doi.org/10.1016/S0140-6736\(17\)31618-5](http://dx.doi.org/10.1016/S0140-6736(17)31618-5).
 12. Machado MA, Maciel AA, de Lemos LLP, Costa JO, Kakehasi AM, Andrade EIG, et al. Adalimumab in rheumatoid arthritis treatment: a systematic review and meta-analysis of randomized clinical trials. *Rev Bras Reumatol*. 2013;53:419-30.
 13. Ginzler EM, Dooley MA, Aranow C, Kim MY, Buyon J, Merrill JT, et al. Mycophenolate mofetil or intravenous cyclophosphamide for lupus nephritis. *N Engl J Med*. 2005;353:2219-28, <http://dx.doi.org/10.1056/NEJMoa043731>.