

RCT Design for the Assessment of Rehabilitation Treatments: The Case Study of Post-stroke Rehabilitation

Simona Ferrante, Marta Gandolla, Elisabetta Peri, Emilia Ambrosini and Alessandra Pedrocchi

1 Introduction

All of us, in our daily lives, perhaps fail to appreciate properly the extraordinary equilibrium that exists between our body and mind, which harmoniously combine to shape who we are. But as soon as we are affected by any disorder that modifies our daily life and habits, we realize just what a gift we had, and we want it back. This is why rehabilitation is so important.

Rehabilitation is the process of helping an individual to regain the highest possible level of functioning, independence and quality of life. The role of neuromotor rehabilitation, in particular, is increasingly recognized. Hence, growing importance is being attached to the evaluation of rehabilitation treatments in the panorama of the available treatment options. Thanks to ongoing technological progress, the scientific and medical communities are now able to propose a wide variety of devices, treatments and integrated solutions for tackling disease. However, for patients, the important questions before starting a specific rehabilitation treatment are always the same: Will I benefit functionally from this treatment? Is it the best treatment I can get? Will I go back to how I was before? Indeed, from a patient's point of view, improvement is not primarily a change in specific parameters, but corresponds,

S. Ferrante (✉) · M. Gandolla · E. Peri · E. Ambrosini · A. Pedrocchi
NearLab, Department of Electronic, Information and Bioengineering,
Politecnico Di Milano, Milan, Italy
e-mail: simona.ferrante@polimi.it

E. Ambrosini
Physical Medicine and Rehabilitation Unit, Scientific Institute of Lissone,
S. Maugeri Foundation, IRCCS, Lissone, Italy

rather, to improvements in functional performance, daily life activities and independence, and they want to be sure that they get the best treatment option available. It is therefore mandatory to be able to answer their questions.

The scientific community recognizes the need for classification of the level of reliability of healthcare information, in order to formulate recommendations for clinical practice. Indeed, all physicians are required to find the highest level of evidence to answer any specific clinical question. The level of evidence is strongly related to the study design, and it is usually represented by the so-called pyramid of evidence (see Fig. 1).

What the pyramid suggests is that not all evidence or information is necessarily equivalent. Systematic reviews of randomized controlled trials (RCTs) are the highest source of stable and generalizable information. They were inspired in the 1980s by Archie Cochrane, who proposed gathering information from works that, if considered singly, failed to provide full conclusions for clinical practice. Often, meta-analysis is used to statistically combine data from several studies, to increase the power and the precision of the single studies' findings. Moving down the pyramid towards the base, we find RCTs. These are widely recognized as the gold standard for drawing statistically acceptable conclusions; they are studies carried out, applying a rigorous methodology, to test the efficacy of a given treatment in a well-defined target population. Other study types, which may be found towards the base of the pyramid of clinical evidence, are accepted in the scientific community, even though the lack of clear and codified methodology often makes the results, although interesting, poorly generalizable.

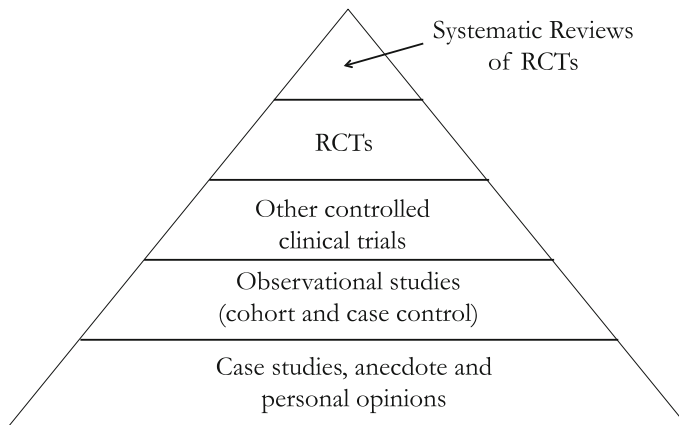


Fig. 1 Pyramid of evidence based on study design: the more systematic the observations the better the quality of the evidence

1.1 What Are RCTs?

Randomized controlled trials are studies carried out, applying a rigorous methodology, to test the efficacy of a given treatment in a well-defined target human population. The key features of an RCT are indicated by its name: the term *randomized* refers to the fact that participants are allocated to one group or another by randomization; *controlled* refers to the fact that the new treatment is given to a group of patients (called the experimental group), while another treatment, often the one currently most widely used, is given to another group of patients (called the control group). RCTs, as we know them today, are relatively recent, dating back to the time of the Second World War, with only isolated attempts made before then [12]. The precision and thoroughness of the methodology are crucially important aspects of RCTs, which, over the past 50 years, have become established as the primary and, in many instances, only acceptable source of evidence for establishing the efficacy of new treatments [12]. As stated by the Oxford Centre for Evidence-Based Medicine, Level I evidence for a given treatment efficacy is assigned if it is supported by multiple RCTs with a statistically significant number of subjects enrolled (www.cebm.net/ocebmllevels-of-evidence). Accordingly, the findings of a well-designed RCT can influence decision-making processes in healthcare, the development of international evidence-based clinical pathway guidelines, and, ultimately, the formulation of national public health policies.

With this in mind, the Consolidated Standards of Reporting Trials (CONSORT) group was created to alleviate the problems arising from inadequate reporting of RCTs (www.consort-statement.org). CONSORT published its first guidelines in 1996 [4], and these were subsequently revised in 2001 [1, 15] and again in 2010 [13]. The scientific community is currently striving to promote convergence among trials, to ensure that they are all well designed and well reported, but there still is a long way to go.

Hopewell and colleagues [9], examined the PubMed-indexed RCTs identified in the database for the years 2000 and 2006 in order to assess whether the quality of reporting improved after publication of the (CONSORT) Statement in 2001. Although they found a general improvement in the presentation of important characteristics and methodological details of RCTs, the quality of reporting was still not considered acceptable. In another study, analyzing only parallel-group design RCTs published in 2001–2010 and in 2011–2014, it was found that 88% and 94% respectively were pharmacological RCTs, meaning that non-pharmacological trials accounted for 12% and 6%; what is more, only 0% and 3% were non-pharmacological trials focusing on the use of devices, meaning that studies dealing with the use of devices had a total incidence of just 1% [18]. With advancing technology now contributing more and more to the development of rehabilitation approaches, there is clearly a need to design and perform rigorous RCTs to clarify the value of new technologies in this setting.

1.2 RCTs—The Need for Randomization

Randomization is the process of allocating study participants to the treatment(s) group(s) or the control group. It is possible that the population of interest could show certain characteristics liable to influence the selected outcome measures, and it is necessary to prevent these characteristics from being concentrated within a given treatment group, as this could result in a systematic effect between the groups that is quite distinct from any treatment effect (i.e., a confounding effect). In the long run, random allocation will equalize individual differences between groups, thus allowing, as far as possible, the treatment effect to be established without being contaminated by any potentially competing factors. In other words, the aim of random allocation is to minimize the effect of possible confounders, leading to a fair comparison between the treatment(s) under investigation and the other procedure chosen as control.

1.3 RCTs—The Need for a Control Group

The efficacy of a novel experimental treatment cannot be demonstrated merely by applying it in a group of patients, because improvements/worsening might also occur spontaneously or following other procedures. In order to demonstrate that the effects observed are attributable exclusively to the treatment under investigation, this needs to be compared with the best treatment currently available. If there is no accepted treatment available, then the control group may receive no treatment at all (which corresponds to what would happen in clinical practice in this situation), or they may receive a sham treatment known as a placebo [12]. It has to be noted that a clear set of eligibility criteria must be defined, and strictly applied, both for the experimental intervention group and for the control group, otherwise it becomes practically impossible to identify the population of patients to which the study results may subsequently be generalized.

1.4 RCTs—The Need for a Statistical Approach

In RCTs, efficacy is not evaluated on a patient-by-patient basis; the aim, rather, is to assess the efficacy of a treatment in a population as a whole. Within each group, there will be patients who perform “outstandingly” and others whose performances will correspond to the “worst” that is observed, as treatments are known to give different outcomes even in a population that is, as far as possible, homogenous.

Since the purpose of RCTs is to demonstrate the superiority (or, less often, the equivalence or non-inferiority) of a novel intervention compared to another, it is necessary to adopt a statistical approach allowing the mean behavior of the target population to be analyzed and reported. This approach requires a well-defined target population and outcome measures, which in turn determine the sample size required in order to draw conclusions, as will be detailed later in the chapter.

1.5 RCT Design

Different study designs are envisaged for RCTs (Fig. 2).

The following five designs, in particular, are commonly implemented:

- parallel-group trials: each single participant is randomized to one of the study arms (i.e., intervention or control group);
- crossover trials: each single participant is exposed to each condition in a random sequence;
- cluster trials: pre-defined homogeneous clusters of individuals (e.g., clinic 1 and clinic 2) are randomly allocated to different study arms (i.e., intervention or control group);
- factorial trials: each single participant is randomly assigned to individual interventions or a combination of interventions (e.g., participant 1 is allocated to intervention x and placebo y; participant 2 is allocated to intervention x and intervention y, participant 3 is allocated to intervention y and placebo y, etc.);
- split-body trials: for each participant, body parts (e.g., upper limbs, lower limbs, etc.) are randomized separately [9].

Among the RCTs indexed in PubMed between 2000 and 2006, 78% of reports were parallel-group trials, 16% crossover trials, and the remaining 6% were classified as “other”; specifically, this last category comprised cluster (33%), factorial (26%), and split-body (41%) trials. Among the parallel-group trials, 76% compared two groups [9]; indeed, the parallel design with two groups is by far the most used, and we therefore concentrate on this particular design in the subsequent sections.

Regardless of the chosen design, RCTs should be carried out as multicenter studies. Indeed, multicenter clinical trials, especially if the different centers are located in ethnically diverse geographical regions, have the main advantage of recruiting a more heterogeneous sample of subjects and involving different care providers. This strengthens the generalizability (external validity) of the investigation [6]. Furthermore, multicenter trials allow an ample sample size to be reached in a reasonable time.

2 Ethical Issues

The peculiarity of an RCT is that the object of investigation is the human being. This aspect differentiates it from all other types of research and makes it necessary to give careful consideration to the ethical principles that regulate the administration, in humans, of interventions with unknown effects.

In 1964, the need for universal ethical rules to govern clinical medical research led the World Medical Association to establish the Declaration of Helsinki, which incorporates ethical principles designed to regulate experiments conducted in humans. Since 1964, this document has been reviewed several times, most recently in 2013 [3]. The Declaration of Helsinki is the ethical benchmark for clinical studies and no reputable medical journal will publish studies whose designs are not based on the principles it sets forth. The main points of the Declaration are summarized below:

1. Experiments must be carried out in compliance with the principle of respect for the patient. The achievement of new knowledge can never supersede the rights of each individual.
2. A treatment that is known to be inferior should not be given to any patient in either the experimental group or the control group.
3. The privacy of the patients must be ensured and their personal information must remain confidential.
4. A patient's participation in the study is voluntary. He/she has the right to refuse to participate, and to withdraw from the study at any time. An informed written consent document, explaining the aim(s) of the study, the possible conflicts of interest, and the possible benefits and the risks entailed in the research should be read and signed by each participant before enrolment.
5. Research can be conducted only if the importance of the objectives is greater than the risks involved. A systematic analysis of the risks must be conducted before the start of the study and every effort should be made to ensure their minimization and to monitor them during the study.
6. A clear research protocol should be established. It should provide the background to the proposal, as well as indications concerning funding, affiliations of researchers, potential conflicts of interest, and compensation for subjects who might be harmed as a result of participation in the research study.

The ethics committee is the body responsible for deciding whether or not a clinical research study should be carried out. To guarantee its transparency and also that it is independent of the interests of any single category (e.g. researchers or sponsors), it is composed of individuals with different roles and areas of expertise, i.e. clinicians, pharmacists, a biostatistician, patients, and experts in medical devices, bioethics, legal and insurance issues. Ethical approval is mandatory and must be obtained before starting any study involving human subjects. Furthermore, the ethics committee has the right to monitor the progress and evolution of the study, and no amendment to the protocol may be made without its consent. At the

end of the study, a report summarizing its main findings should be submitted to and evaluated by the ethics committee.

3 Bias in RCTs

Bias in clinical trials is defined as a systematic error that may induce misleading conclusions about the efficacy of one treatment over others. When designing an RCT, it is necessary to avoid all possible sources of systematic errors that could affect the treatment effect estimation. Indeed, the quality of clinical trials can be measured by evaluating the robustness of the design in this regard. However, in scientific reports, little effort is made to discuss methods for addressing sources of potential bias [10]. Typical biases in clinical trials can be classified in the following five categories [12]:

1. *Selection bias*

This type of bias can occur during the selection of the patients to be enrolled in the study. Ideally, each patient should be enrolled in the study before being randomized and allocated to a treatment group (Fig. 2). Indeed, were the medical doctor to know, in advance, the treatment group to which patients have been allocated, this could influence his/her decision to enrol a given patient.

2. *Allocation bias*

When treatment groups are not balanced in size and not similar at baseline, allocation bias is encountered [12]. The randomization procedures (Fig. 2) should be designed to balance groups at baseline with particular attention to the prognostic factors that could influence the outcome (e.g., age-matched groups).

3. *Assessment bias*

When the assessor is not blinded to treatment allocation, his/her evaluation could be influenced, leading to assessment bias (outcome measures assessment, Fig. 2). This problem is particularly important in the case of outcome measures that are characterized by low sensitivity, high subjectivity (such as patient-reported outcome measures), and low inter- and intra-rater reliability. Moreover, assessment bias can also occur when the patient is not blinded to treatment allocation, since this may lead him/her to behave differently during the assessment [12]. A recent review highlighted that trials with inadequate allocation concealment or lack of blinding tend to overestimate the intervention effect, when compared with adequately concealed trials [21]. The recent use of technologically advanced tools to flank traditional assessments based on functional scales may be helpful to reduce assessment bias. Indeed, technology provides quantitative, non-operator-dependent evaluations and promotes a consistent comparison among results of different clinical trials. Nevertheless, it often requires expensive tools that are not always available in clinical practice.

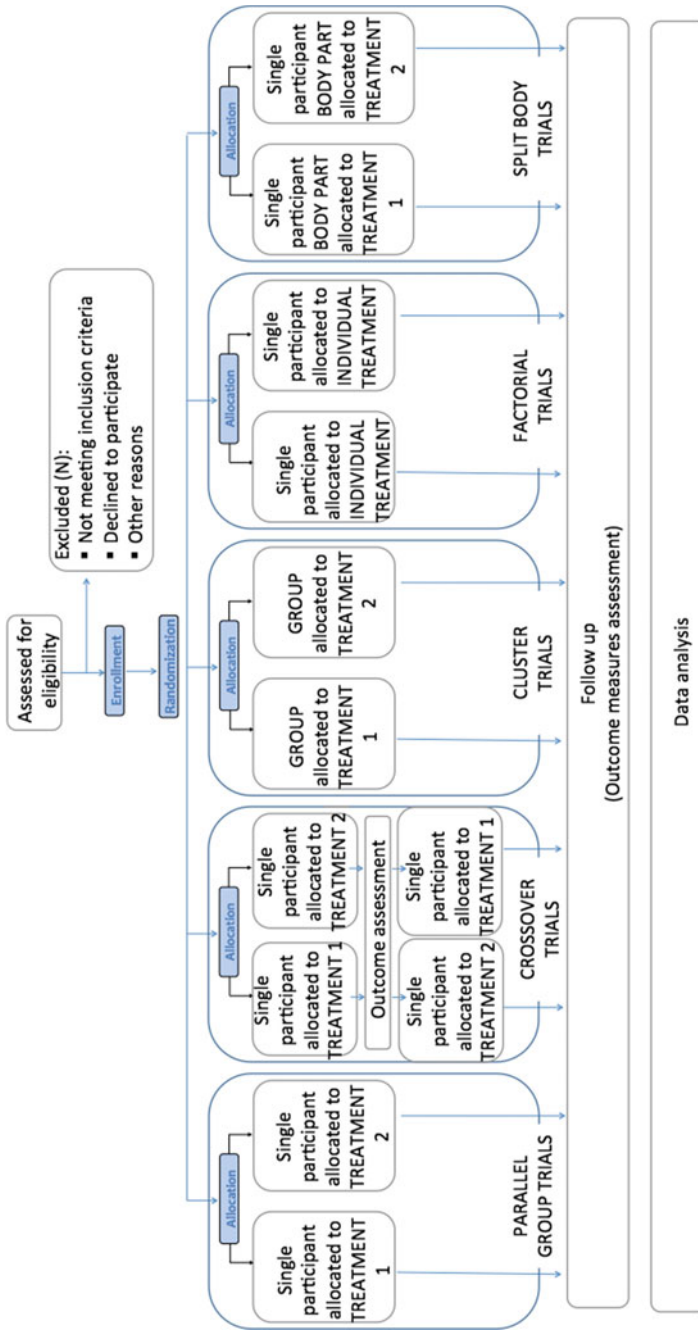


Fig. 2 Flowchart of the different RCT designs

4. *Publication bias*

The ultimate goal of any medical research is to influence clinical practice. A fundamental step, in order to share RCT findings, is to have a report of the trial published in an international scientific journal. To be published, papers undergo a peer review process, in which several experts are called upon to evaluate the quality of the clinical trial itself, and also of the report. Papers reporting positive findings are often considered more likely to be published than papers that do not show any statistically significant difference [12]. This leads to a clear bias, with the medical literature not giving equal exposure to studies with positive and negative results.

5. *Stopping bias*

To avoid a stopping bias, it is necessary to make an a priori decision on exactly when the recruitment of patients will be stopped. There are two possible choices: the first is to keep on recruiting until it is established which treatment is superior. The second is to decide beforehand the number of patients that should be recruited for the RCT. This second method is the one usually preferred, because of the bias that can potentially be induced by stopping a trial when the results obtained are considered satisfactory. Deciding the sample size in advance is not a trivial aspect. Indeed, recruiting too few patients may prevent the achievement of definitive results on the superiority of a given treatment over others. On the other hand, recruiting more patients than necessary raises an ethical issue, since the surplus patients could all be exposed directly to the superior treatment.

In response to growing evidence of bias resulting in misleading clinical trial results, several bodies have, in recent years, tried to develop guidelines to guarantee the quality of clinical trial reporting [11]. Among these, the CONSORT group defined a 25-point checklist intended to promote rigorous reporting of research and, indirectly, to prevent the publication of inadequately designed studies [14]. The checklist is intrinsically developed to minimize the risk of bias.

4 Guidelines on Designing a Parallel-Group RCT

RCTs should be adequately designed and transparently reported in order to allow critical appraisal of the validity and applicability of their results. To assess the validity of novel interventions in rehabilitation, a field in which blindness and reproducibility of the treatment are not easy to achieve, the best guidelines to follow are the CONSORT guidelines for non-pharmacological studies [19]. Below, we summarize the guidelines that should be followed in the design of a parallel-group RCT in the rehabilitation field. For each step, the example of a currently ongoing case study is provided to better clarify the point made [8].

4.1 Trial Design

The organization of an RCT starts with the formulation of a clinical question (hypothesis) that derives from the observation of patients. To demonstrate the correctness of the hypothesis a rigorous study design should be defined a priori, establishing not only the trial type (parallel group, crossover, etc.), but also the conceptual framework. In parallel-group RCTs, the conceptual framework corresponds to the hypothesis which the trial aims to demonstrate (i.e., the superiority, equivalence or non-inferiority of the novel intervention with respect to the other).

Case study. The study proposes a novel rehabilitation program for the recovery of locomotor abilities in post-acute stroke patients. The conceptual framework is that a specific rehabilitation program (involving biofeedback cycling training, combining voluntary effort and Functional Electrical Stimulation (FES) of the leg muscles, and biofeedback balance training) is superior to standard therapy in improving walking abilities, disability, motor performance, and independence in post-acute stroke patients.

In view of this conceptual framework, the study was designed as a parallel-group (two-group) superiority RCT.

4.2 Ethical Approval and Trial Registration

To avoid carrying out studies whose value is undermined by methodological weaknesses, lack of transparency and/or conflicts of interest, the RCT's principal investigator prepares, as requested by an ethics committee, a series of detailed documents describing the study protocol. The RCT should be approved by this committee before the study is started.

Moreover, as already mentioned, both authors and journal editors prefer studies that give positive findings, and this leads to a publication bias, i.e. an over-representation of positive studies. One method for reducing publication bias is to register all studies undertaken in a recognized international repository. On registering the RCT, the principal investigator should declare and describe the complete methodology. Indeed, an RCT should be registered soon after securing ethical approval and before the enrolment of the first patient. Nowadays, registration is becoming mandatory for publication in the main rehabilitation journals.

Case study. The study received the approval of the Central Ethics Committee of the rehabilitation center in which it is being carried out (Comitato Etico Centrale of the Salvatore Maugeri Foundation). The trial has been registered at clinicaltrials.gov (Identifier: NCT02439515).

4.3 *Participants' Eligibility*

The profile of patients who may benefit from the intervention under investigation should be defined through clear inclusion/exclusion criteria. The more explicit these criteria are, the more the replicability of the results will be facilitated. Indeed, the aim of the trial is to obtain results that can be generalized and applied to all patients similar to those treated in the trial.

Normally the number of eligible patients is different from the number of recruited patients as a result of the need to apply the necessary ethical rules before enrolment. According to these rules, each patient should be informed in detail about the trial and should decide voluntarily whether or not to participate. Patients deciding to participate should then sign an informed consent document.

In a rehabilitation RCT it is very important to define the participants' eligibility, but also the rehabilitation centers in which the trial is performed. Indeed, the level of expertise of the care providers should be clearly stated, in order to ensure that others are equipped to reproduce the treatment correctly.

Case study: The study participants are adult post-acute stroke patients who experienced a first stroke less than 6 months before recruitment. Other inclusion criteria were a low level of spasticity of the leg muscles (Modified Ashworth Scale score < 2), no limitations at the hip, knee, and ankle joints, and the ability to sit comfortably and independently for 30 min. Given that the experimental intervention includes FES, people with pacemakers and/or allergic to adhesive stimulation electrodes were excluded.

4.4 *Outcome Measures*

In all rehabilitation RCTs, changes in the patients' conditions induced by the treatment(s) under investigation are assessed by collecting so-called outcome measures. These consist of internationally recognized clinical tools such as clinical scales (questionnaires or rating scales administered by the rehabilitation staff, who ask patients to perform cognitive or motor tasks) and technology-based assessment, e.g. biomechanical tests to obtain quantitative data for movements involved in the interventions, diagnostic examinations, and so forth.

When designing an RCT it is necessary to select a primary outcome measure that is strictly related to the final aim of the intervention. This measure plays a crucial role, as it is the element taken into account by the principal investigator in order to decide how large the trial ought to be. It is preferable to choose, as the primary outcome, a well-recognized measure for which it is possible to find, in the literature, the minimum clinically important difference in the population under investigation. In addition to the primary outcome, all the other measures of interest, i.e., the secondary outcome measures, should be defined a priori.

When applicable, any methods that could enhance the quality of measurements (e.g., multiple observations, training of assessors) should be defined and used.

It is also crucial to define precise time points at which all the outcome measures are collected. They should usually be collected before the beginning of the intervention, soon after the end of the intervention, and in a follow-up session, which should take place at least 6 months after the end of the intervention, thereby making it possible to evaluate changes in the medium and long term.

Case study. The primary outcome measure was gait speed computed during a 10 m walking test performed using GAITRite® (CIR System Inc., USA), a pressure sensor system that allows measurement of spatiotemporal gait parameters. Secondary outcome measures were: other spatiotemporal gait parameters, such as cadence, stride and step length, single and double support time, swing velocity, etc.; clinical scales, such as the Fall Efficacy Scale, the Berg Balance Scale and the Global Perceived Effect scale; the distance covered during the 6 min walking test; lower limb EMG activations and pedal forces produced while pedaling on a sensorized cycle ergometer; and measures of postural control obtained using a sensorized balance board.

The outcome measures were collected at baseline, soon after the end of the intervention and at 6 month follow-up.

4.5 Intervention Programs

When designing an RCT a precise definition of the interventions under investigation (both the novel one and the one used as control) should be provided. It is very important to ensure that, overall, the amount of training received by the groups is balanced.

Case study. The experimental program consisted of 15 sessions of FES-supported volitional cycling training followed by 15 sessions of balance training. Both cycling and balance training were supported by visual biofeedback in order to maximize patients' involvement in the exercise and were performed in addition to standard therapy. The control group received standard physical therapy consisting of stretching, muscle conditioning, exercises for trunk control, standing, and walking training, and upper limb rehabilitation. Both training programs lasted 6 weeks and patients were trained for about 90 min daily. Cycling and balance training lasted about 20 min, therefore the patients in the experimental group performed only about 70 min of standard therapy. This choice ensured that the overall duration of the intervention was the same in both groups.

4.6 Sample Size

Under the hypothesis of normal distribution of the primary outcome measure, the equation used to determine the sample size for a parallel-group RCT with two groups is the following:

$$N = \frac{2\sigma^2 \left(z_\beta + z_{\frac{1-\alpha}{2}} \right)^2}{\tau_M^2}$$

where τ_M is the treatment effect, i.e., the between-group difference under the hypothesis that the baseline of the two groups is similar; σ is the standard deviation of the outcome measure in the population under investigation; z_β and $z_{\frac{1-\alpha}{2}}$ are the quantiles of the standard normal distribution in which β is the false positive rate, i.e., a measure of the desired power (usually set to 90% or 80%), and α is the significance level (usually set at 0.05 or 0.01). When available, τ_M is defined as the minimum clinically important difference of the outcome measure in the population under investigation.

Thus, according to this equation, the “noisier” the measurements are, i.e., the larger σ is, the larger the sample size needs to be. Also, a larger sample size is needed in order to detect a smaller difference in the means obtained in the two treatments.

Once the sample size needed in order to be able to observe a treatment effect with a given statistical power has been determined, the number should be increased to account for the possible dropout rate (normally a rate of 20% is acceptable).

Case study. Gait speed was chosen as the primary outcome measure. Sample size was defined on the basis of ability to detect a minimal clinically important difference for gait speed, estimated as 0.16 m/s (i.e., $\tau_M = 0.16$ m/s) with a standard deviation (σ) of 0.22 m/s. Given that a power of 80% and a significance level of 0.05 were chosen, a sample of 60 subjects was obtained. Thus, allowing for a 20% drop-out rate, 72 patients had to be recruited.

4.7 Randomization and Allocation Concealment

A good randomization procedure should include the generation of an unpredictable allocation sequence and the concealment of that sequence until assignment occurs [17]. Different methods have been proposed:

Simple randomization. This method requires the generation of a sequence of “A” and “B” values. Each entry in the sequence is independent of all other elements and equally likely to take both possible values. This is equivalent to the generation of a sequence in which each entry is defined by the toss of a coin. This method can

result in an unbalanced number of patients per group or an unwanted imbalance between the groups at baseline.

Restricted randomization. Two main methods can be implemented: blocking randomization, which ensures that groups will be of approximately the same size, and stratification randomization, which ensures that groups will be comparable at baseline.

For blocking randomization, it is recommended to use the random permuted blocks method, based on the use of blocks of random size, because it completely avoids any selection bias. This randomization requires the choice of two even-number block sizes and it defines all the permutations of balanced values possible for each of them (e.g., if the block size is 2 there can be only two permutations, AB and BA), but if blocks of 4 and 6 elements are chosen, for example, we will have a list of 6 sequences of 4-element blocks and a list of 20 sequences of 6-element blocks (for a block of size N , the number of all possible permutations of elements is equal to $N!/((N/2)!(N/2)!)$). Then, at each iteration of the sequence generation, both the size of the block (i.e., 4 or 6) and the number of the sequence in the list (i.e., a number between 1 and 6 when 4 is extracted and a number between 1 and 20 when 6 is extracted) are randomized to create the final allocation list.

In the stratification method, the randomization can be simple or by blocks, but it is carried out separately within each of two or more subsets of participants (for example, corresponding to disease severity or study centers) in order to ensure that the patient characteristics are closely balanced within each intervention group.

In rehabilitation RCTs, it is not only the patients recruited that are randomized, but also the clinicians who will administer the intervention. Indeed, often the two interventions require different levels of clinical expertise. Therefore, the best practice to avoid biased results and improve the applicability and generalizability of the findings is to train all the possible clinicians to the same level, and then select a random sample from among them [5]. The situation is more complex, however, when the additional expertise required is not just training in the use of a novel device. Indeed, in such circumstances (e.g. surgery trials), it is impossible to recruit only clinicians who are able to perform both treatments. Recommendations for these special cases are reported in Devereaux et al. [7] and these RCTs are defined expertise-based RCTs.

Case study. Subjects were randomized to one of two groups, one undergoing the novel rehabilitation program in addition to standard therapy (experimental group), and one undergoing standard therapy alone (control group). The randomization criteria applied were blocking randomization using blocks of random size as explained above. Given the impossibility of predicting the next value in the randomization sequence, this randomization method ensured allocation concealment.

4.8 *Blinding*

Complete blindness of an RCT means that the participants, those administering the interventions, and those assessing the outcomes have to be completely unaware of the group assignment. In some studies, blindness of patients and care providers is impossible; obvious examples are trials comparing surgical with nonsurgical treatment, or different rehabilitation interventions to demonstrate the superiority of a novel device. In these cases, the minimum condition required is that the assessors be blinded to the group assignment.

A good but very difficult way of obtaining participant blindness is to compare a novel intervention with a placebo version of the same intervention. The placebo should look identical to the active treatment. One attempt to achieve this in the rehabilitation field was made by Ambrosini et al. in a study that compared an FES cycling treatment with an FES cycling placebo [2]. The patients included in the placebo group had electrodes attached correctly to their skin and underwent the same cycling sessions as the other group, but no stimulation current was delivered to their muscles. An easier and more common use of placebo concerns drug use. An example in post-stroke rehabilitation is described by Ward and colleagues (2017), who treated a chronic stroke population with motor training associated with either pharmacological modulation of neuroplasticity (atomoxetine) or placebo drugs [20].

4.9 *Statistics*

The method of statistical analysis should be chosen a priori during the study design phase. To analyze the treatment effect, a linear mixed model able to handle covariates (in the event of baseline imbalances) and missing values is recommended. When a significant difference between the two groups is found, it is crucial to understand whether this difference is also clinically relevant, in other words, greater than the minimum clinically important difference for the specific outcome measure in the population under investigation.

Case study. Given the randomization method chosen (blocking randomization), the statistical analysis, performed at the end of the study, needed to take into consideration, among other things, the possibility that there was some imbalance in the baseline characteristics. For this reason, the principal investigator decided to use a linear mixed model in which, if necessary, covariates can be considered.

5 Conclusions

This “guided tour” of the process of RCT design has highlighted the ethical and methodological principles that must be adhered to in order to demonstrate, correctly, the efficacy of a given treatment in the rehabilitation field.

Although RCTs are considered a sound basis for demonstrating treatment efficacy, they also present some limitations. First, a properly designed and conducted RCT implies quite significant costs, and therefore usually needs to be supported by a dedicated research grant. Second, many outcome measures of interest can only be verified some considerable time in the future. We may think, for example, of the long-term effects of dietary modifications on stroke prevention, or of fall prevention interventions designed for implementation in the healthy aging population—their results may not be available for decades. This makes it difficult to follow up enrolled patients effectively, and also makes the costs substantially greater than those (for personnel and resources) associated with the conducting of the trial itself. This may explain why so many RCTs encounter difficulties—failing to achieve the sample size needed in order to observe, and therefore report, the desired effect [16]. Finally, Level I evidence for a given treatment efficacy is assigned if it has been the focus of multiple RCTs (or multicenter studies). But replicating RCTs, i.e. focusing on a treatment already investigated, has two immediate implications: (i) difficulty obtaining funding, due to (ii) the lack of innovation in the study proposal.

Randomized controlled trials, even though their value is often underestimated, are the real bedrock on which to build effective conclusions on new rehabilitation treatments and can impact on clinical practice. Indeed, a well-designed RCT can influence decision-making processes in healthcare and the international evidence-based clinical pathway guidelines, and ultimately, benefit patients, who should always be placed at the heart of RCTs, and of rehabilitation research in general.

Acknowledgements This research was supported by the Italian Ministry of Health (grant no.: GR-2010-2312228, title “Fall prevention and locomotion recovery in post-stroke patients: a multi-modal training towards a more autonomous daily life”).

References

1. Altman DG, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med.* 2001;134(8):663–94.
2. Ambrosini E, et al. Cycling induced by electrical stimulation improves motor recovery in postacute hemiparetic patients: a randomized controlled trial. *Stroke; J Cereb Circ.* 2011;42(4):1068–73.
3. Anon. World medical association declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA.* 2013;310(20):2191–4.
4. Begg C, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA.* 1996;276(8):637–9.

5. Boutron I. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med.* 2008;148(4):295.
6. Chung KC, Song JW, WRIST Study Group. A guide to organizing a multicenter clinical trial. *Plast Reconstr Surg.* 2010;126(2):515–23.
7. Devereaux PJ, et al. Need for expertise based randomised controlled trials. *BMJ: Br Med J.* 2005;330(7482):88.
8. Ferrante S. et al. Does a multimodal biofeedback training improve motor recovery in stroke patients ? Preliminary results of a randomized control trial. 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (*EMBC 2015*). 2005;42, p. 2011.
9. Hopewell S, et al. The quality of reports of randomised trials in 2000 and 2006: comparative study of articles indexed in PubMed. *BMJ.* 2010;340:c723.
10. Horton R. The rhetoric of research. *BMJ (Clinical research ed.)*. 1995;310(6985):985–7.
11. Johansen M, Thomsen SF. Guidelines for reporting medical research: a critical appraisal. *Int Sch Res Not.* 2016; vol. 2016:1346026, p. 7.
12. Matthews JNS. *Introduction to Randomized Controlled Clinical Trials*. Second Edition, 2006.
13. Moher D, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ (Clinical research ed.)*. 2010;340:c869.
14. Moher D, et al. CONSORT 2010 explanation and elaboration: Updated guidelines for reporting parallel group randomised trials. *Int J Surg.* 2012;10(1):28–55.
15. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *J Am Podiatr Med Assoc.* 2001;91(8):437–42.
16. Sanson-Fisher RW, et al. Limitations of the randomized controlled trial in evaluating population-based health interventions. *Am J Prev Med.* 2007;33(2):155–61.
17. Schulz KF. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *JAMA: J Am Med Assoc.* 1994;272(2):125.
18. Stevely A, et al. An Investigation of the shortcomings of the CONSORT 2010 statement for the reporting of group sequential randomised controlled trials: a methodological systematic review. *PLoS One.* 2015;10(11):e0141104.
19. Villamar MF, et al. The reporting of blinding in physical medicine and rehabilitation randomized controlled trials: a systematic review. *J Rehabil Med.* 2013;45(1):6–13.
20. Ward A, et al. Safety and improvement of movement function after stroke with atomoxetine: a pilot randomized trial. *Restor Neurol Neurosci.* 2017;35(1):1–10.
21. Wood L, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ (Clinical research ed.)*. 2008;336(7644):601–5.