RBFEx
ISSN Online: 2675-1372

Revista Brasileira de Fisiologia do Exercício

Opinion

Methodological challenges in randomized clinical trials in physical education: the design of non-inferiority

Desafios metodológicos em ensaios clínicos randomizados na educação física: o design de não inferioridade

Antônio Marcos Andrade Costa[1,2,3], Ewerton de Souza Bezerra[2], Nathalia Bernardes[3]

1. Universidade Estadual do Amazonas (UEA), Manaus, AM, Brazil
2. Universidade Federal do Amazonas (UFAM), Manaus, AM, Brazil
3. Universidade São Judas Tadeu, São Paulo, SP, Brazil

ABSTRACT

The Physical Education professional, like any health professional, needs to make decisions during the exercise of his professional activity. These decisions must be prudent, aiming for the greatest benefit for your client. In this context, randomized clinical trials (RCTs) are considered the gold standard to guide decision making. In this context, randomized clinical trials (RCTs) are considered the gold standard to guide decisions. However, mistaken judgments can occur when interpreting the results of clinical superiority studies, because they assume that two interventions are identical due to the absence of statistical difference, however, the lack of statistical significance does not support the conclusion of equality; that is, the absence of evidence is not evidence of absence. In this scenario, an elegant alternative is equivalence and non-inferiority studies, which should be used whenever a new intervention has a substantial practical advantage compared to the old, already established one. According to the methodological strategy, a tolerance margin for non-inferiority is established using the limits of the confidence interval. In this way, once non-inferiority has been demonstrated, we become more convinced that the intervention will bring the expected benefit to our client. Therefore, our proposal was to draw attention to this methodological technique that can be of great use in our area and that needs to be further explored.

Keywords: physical education; randomized clinical trial; non-inferiority.

RESUMO

O profissional de Educação Física, como qualquer profissional de saúde, necessita tomar decisões durante o exercício da sua atividade profissional. Essas decisões devem ser prudentes visando o maior benefício para o seu cliente. Neste contexto, os ensaios clínicos randomizados (ECR) são considerados o padrão ouro para orientar a decisão. No entanto, julgamentos equivocados podem acontecer na interpretação dos resultados de estudos clínicos de superioridade, isto porque assumem que duas intervenções são idênticas devido a ausência de diferença estatística, todavia, a falta de significância estatística não apoia a conclusão da igualdade; isto é, a ausência de evidência não é evidência de ausência. Neste cenário, uma alternativa elegante são os estudos de equivalência e não inferioridade, que devem ser utilizados sempre que uma nova intervenção tenha uma vantagem prática substancial em comparação com a antiga já estabelecida. De acordo com a estratégia metodológica, é estabelecida uma margem de tolerância para não inferioridade utilizando os limites do intervalo de confiança. Dessa forma, uma vez demonstrada a não inferioridade, ficamos mais convencidos que a intervenção trará benefício esperado para nosso cliente. Portanto, nossa proposta foi chamar a atenção para essa técnica metodológica que pode ser de grande utilidade em nossa área e que necessita ser mais explorada.

Palavras-chave: educação física; ensaios clínicos randomizados; não-inferioridade.

Received: June 10, 2024; Accepted: July 18, 2024.

Correspondence: Antonio Marcos Andrade Costa, antoniomarcoshand@gmail.com

Introduction

The Physical Education professional, like any other health professional, needs to make decisions during the exercise of their clinical activity. These decisions must be prudent and most likely to benefit your patient client. To achieve this, the mental process of judgment that precedes your actions must be based on logical analysis that follows a mental trigger, taking into account your professional expertise, the patient and the evidence regarding the conduct you intend to take. Well-planned and executed clinical trials are the best methodological designs for testing the cause and effect relationship between a set of independent and dependent variables in experimental models [1].

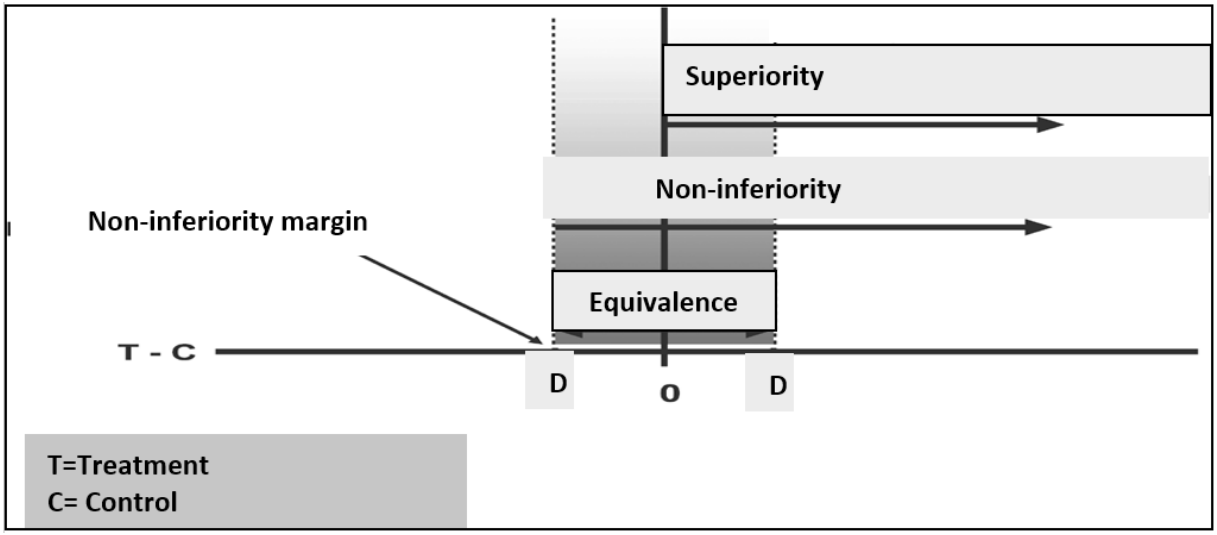
When it comes to interventions involving human subjects, randomized controlled trials (RCTs) are considered by proponents of evidence-based healthcare to be the gold standard design to guide decision-making [2]. Classically, sample groups are defined through random allocation, with one being an experimental group (representing the intervention being tested) and another group being considered the control – which can sometimes be no treatment, a placebo or, more frequently, a recognized efficacy treatment. The results undergo appropriate statistical analysis in order to validate the conclusions and identify the best interventions. This methodological model is called effectiveness superiority study or comparative effectiveness, and its analysis for decision making involves testing hypotheses; the null hypothesis we call H0, and the alternative hypothesis called H1. In this type of experiment, the randomization process, when carried out satisfactorily, makes the groups homogeneous and therefore comparable, eliminating confounding factors, which leads the investigator to reject H0 in the presence of a p value <0.05 , and conclude that the difference observed between the groups comes from the intervention applied.

Frequentist statistics teaches us that, by not rejecting the null hypothesis, we may be facing what we call a type II error – failing to show a relevant difference due to the lack of appropriate sizing at the time of designing the study. This normally occurs under conditions of low statistical power, either due to an insufficient number of participants or due to biases in the design and/or conduct of the study [3]. In superiority study designs, the null hypothesis (H0) states that the intervention tested is not superior to the control group, while the alternative hypothesis (H1) states that the intervention is superior to the control group.

However, we observe misinterpretations when it is not possible to reject the null hypothesis, as it is not uncommon for researchers to conclude that in the absence of statistical difference between interventions, they are equal. Many authors report their results in a way that leads readers to conclude that the interventions "are equivalent", one way to identify this practice is when faced with a finding in which it was not possible to identify a difference, the authors begin to base their narratives solely on biological plausibility, inducing the reader to extract a positive result in the absence of significance [4]. However, the lack of statistical significance does not

support the conclusion of equality between interventions; that is, 'the absence of evidence is not evidence of absence [5]. Inference errors like these have been appearing more frequently and can contribute to the formation of a flawed and unreliable ecosystem.

It is in this scenario that an elegant alternative emerges to test a promising idea that presents a clear practical advantage (low cost, lower risk, low application complexity, among others) which are non-inferiority projects. In this construct, the null hypothesis (H0) states that the new intervention tested is not inferior to the control (old), therefore similar, or that through a plausible a priori argument, it is accepted until the new intervention is less effective (it is established a non-inferiority limit), so that (H0) and the conclusion of non-inferiority can be accepted. In an attempt to materialize a concrete example, below is figure 1, taken from the study "Non-inferiority clinical trials: concepts and issues" [6].



T = treatment; C = control. T is superior to C if the confidence interval of the difference lies entirely to the right of zero, non-inferior if entirely to the right of -D, and equivalent if contained within the equivalence zone between -D and +D

Figure 1 - Theoretical basis for concepts applied during randomized controlled trial

An ideal moment to resort to non-inferiority studies is when a new intervention emerges or when testing something new is necessary. However, the candidate must offer an explicit advantage over the intervention already established in the literature, justifying the acceptance of non-inferiority. An example of an intervention that, in our understanding, would justify testing non-inferiority is high-intensity interval training (HIIT) in the outcome of improving VO2max. Once demonstrated in a non-inferiority design that HIIT can achieve the non-inferiority threshold, we can conclude that the intervention is indeed time-efficient.

In table I, it was reprinted from the study by Pinto [6], in 2010, in which the author presents an analysis algorithm for three types of hypothetical studies, in which T represents the measure of effectiveness of the new intervention, and C the measure of effectiveness of the control group. Rejecting the null hypothesis means, for superiority studies, that the new intervention called T is superior to the control group C;

non-inferiority studies, when the difference between C and T is smaller than a margin delta (D) non-inferiority margin and, for equivalence studies, that the difference between C and T is neither smaller nor larger than a margin D. Fundamentally, the term equivalent means non-inferior and non-superior, and testing for equivalence refers to the analysis for the symmetric region defined by [+D,-D].

Table I - Formulation of hypotheses for superiority, non-inferiority, and equivalence studies

Study desing	Null hypothesis	Alternative hypothesis
Superiority	Ho:C–T≥0	Ha:C–T<0
Non-inferiority	Ho:C – T ≥D	Ha:C–T<D
Equivalence	Ho:|C–T|≥D	Ha:|C– T|<D

Effectiveness measures are presented in the table above as; T-new intervention, C- control, and D as the margin of non-inferiority/equivalence

Moreover, several factors must be carefully considered when planning, analyzing, and interpreting non-inferiority studies to ensure the study's internal validity: a) choice of the non-inferiority margin; b) the number of participants required for the study; c) control of the study's sensitivity; d) definition of the analysis population. Some other factors should be considered as well, but they are beyond the scope of our discussion at the moment.

Conclusion

In the field of research in Physical Education, as well as in other segments of the health sector, we need to pay attention to make the best decisions, and knowing how to interpret the results of scientific findings is an elementary skill, essential to becoming better professionals. Our intention was to alert the academic community and science consumers regarding the conclusion of similarity drawn from the results of clinical trials of superiority of efficacy without statistical difference. As we frequently observe, many professionals justify the application of an intervention based on the inferences drawn from this mental model, which makes the debate relevant and necessary.

We highlight non-inferiority studies as an alternative to address this issue, as we believe it is a more elegant methodological technique, as it establishes specific parameters to test similarity or a limit to accept non-inferiority. For those who wish to delve deeper into the topic, we recommend two reference materials that were valuable in preparing this reflection: one developed by the CONSORT [7] group and the other provided by the Canadian Cancer Partnership at McMaster University [8]. Bringing this topic into a broader discussion and encouraging the development of specific guidelines and guidelines on the subject appears to be an emerging need

Funding
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest statement
The authors have no conflict of interests to declare.

References

1. Portney LG. Experimental Designs. In: Foundations of clinical research: Application to evidence--based practice. 4.ed. Philadelphia, USA: FA Davis Company; 2020.

2. Zabor EC, Kaizer AM, Hobbs BP. Randomized Controlled Trials. Chest. 2020;158(1):S79-87. doi: 10.1016/j.chest.2020.03.013

3. Baldissera R. Uma iniciação aos testes estatísticos para dados biológicos: inclui scripts dos testes em R. Chapecó: UFFS; 2022.

4. Boutron I, Dutton S, Ravaud P, Altman DG. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. JAMA. 2010;303(20):2058–64. doi: 10.1001/jama.2010.651

5. Altman DG, Bland JM. Absence of evidence is not evidence of absence. BMJ 1995;311:485

6. Pinto VF. Ensaios clínicos de não inferioridade: conceitos e questões. Jornal Vascular Brasileiro. 2010;145-51.

7. Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ; CONSORT Group. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. JAMA. 2006;295:1152-60.

8. Canadian Partnership Against Cancer. Capacity Enhancement Program: non-inferiority and equivalence trials checklist [Internet]. McMaster University; 2009. [citado 2024 set 13]. http://fhs.mcmaster.ca/cep/documents/CEPNon-inferiorityandEquivalenceTrialChecklist.pdf