

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

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

Antônio Marcos Andrade Costa<sup>1,2,3</sup> , Ewerton de Souza Bezerra<sup>2</sup> , Nathalia Bernardes<sup>3</sup> 

1. Universidade Estadual do Amazonas (UEA), Manaus, AM, Brasil

2. Universidade Federal do Amazonas (UFAM), Manaus, AM, Brasil

3. Universidade São Judas Tadeu, São Paulo, SP, Brasil

#### 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.

#### 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.

## Introdução

O profissional de Educação Física, assim como outro profissional da saúde, necessita tomar decisões durante o exercício da sua atividade clínica. Essas decisões devem ser prudentes e com maior probabilidade de beneficiar seu cliente. Para isso, o processo de julgamento que antecede suas ações deve ser baseado em análise lógica, levando em consideração sua expertise profissional, o paciente e as evidências a respeito da conduta que pretende tomar. Ensaios clínicos bem planejados e executados são os melhores desenhos metodológicos para testar a relação de causa e efeito entre um conjunto de variáveis independentes e dependentes em modelos experimentais [1].

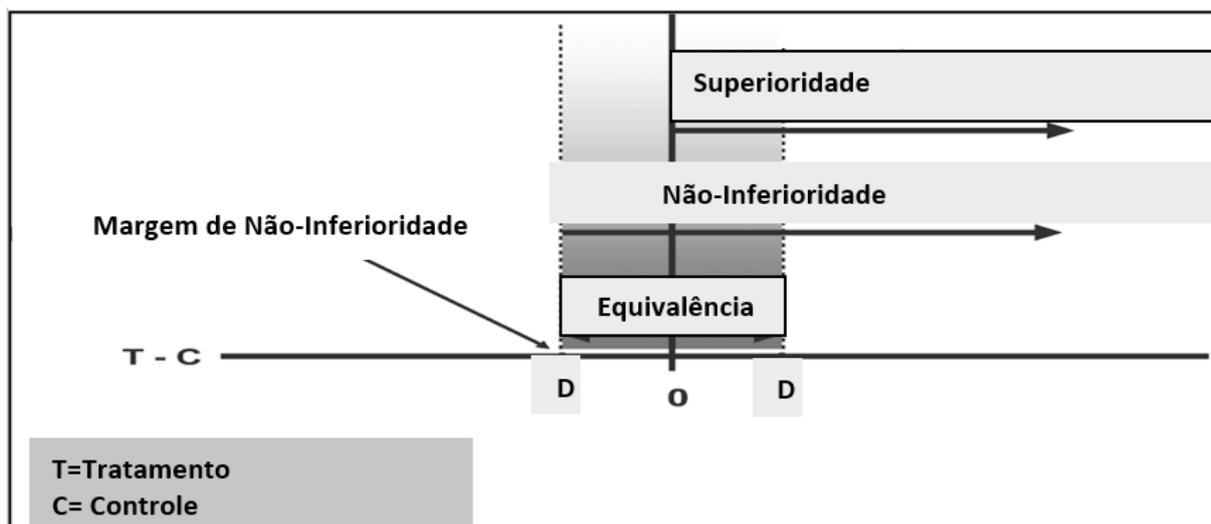
Isso porque os ensaios clínicos randomizados (ECR) são considerados pelos proponentes da saúde baseada em evidências, como o desenho padrão ouro para orientar a tomada de decisões [2]. Classicamente, os grupos amostrais são definidos através de alocação aleatória, sendo um grupo experimental (representando a intervenção que está sendo testada) e outro grupo é considerado o controle que pode ser ausência de tratamento, um placebo ou tratamento de eficácia reconhecida. Os resultados passam por análise estatística adequada a fim de validar as conclusões e identificar as melhores intervenções. Esse modelo metodológico é chamado de estudo de superioridade de eficácia ou eficácia comparativa, e sua análise para tomada de decisão envolve o teste de hipóteses. A hipótese nula que chamamos de  $H_0$ , e a hipótese alternativa denominada  $H_1$ . Neste tipo de experimento, o processo de randomização, quando realizado de maneira satisfatória, torna os grupos homogêneos e assim sendo, comparáveis, eliminando fatores de confusão, o que leva o investigador a rejeitar  $H_0$  na presença de um valor de  $p < 0,05$ , e concluir que há diferenças entre os grupos intervenção e controle.

A estatística frequentista nos ensina que, ao não rejeitar a hipótese nula, podemos estar diante do que chamamos de erro tipo II, que deixa de mostrar uma diferença relevante devido a falta de um dimensionamento apropriado no momento da concepção do estudo. Isso normalmente ocorre em condições de baixo poder estatístico, seja por número insuficiente de participantes ou por viés no desenho e/ou condução do estudo [3]. Em desenhos de estudos de superioridade, a hipótese nula ( $H_0$ ) afirma que a intervenção testada é não superior ao grupo controle, já hipótese alternativa ( $H_1$ ) afirma que a intervenção é superior ao grupo controle.

Contudo, observamos interpretações equivocadas quando não é possível rejeitar a hipótese nula, pois não é incomum que investigadores concluam que na ausência de diferença estatística entre as intervenções, elas sejam iguais. Muitos autores relatam seus resultados de maneira a levar os leitores a concluir que as intervenções “são equivalentes”, uma maneira de identificar essa prática é quando diante de um achado que não foi possível identificar diferença, os autores começam a basear suas narrativas apenas em plausibilidade biológica, induzindo o leitor a extrair um resultado positivo na ausência de significância [4]. Todavia, a falta de significância

estatística não apoia a conclusão da igualdade entre as intervenções; isto é, a ausência de evidência não é evidência de ausência [5]. Erros de inferência como esses têm aparecido com mais frequência e podem contribuir para a formação de um ecossistema falho e pouco confiável.

É neste cenário que surge uma alternativa elegante para testar uma ideia promissora que apresenta uma clara vantagem prática (baixo custo, menor risco, baixa complexidade de aplicação, entre outras) que são os projetos de não inferioridade. Neste constructo, a hipótese nula ( $H_0$ ) afirma que a nova intervenção testada não é inferior ao controle (antiga), portanto semelhante, ou que através de um argumento a priori plausível, ela é aceita até que a nova intervenção seja menos eficaz (é estabelecido um limite de não inferioridade), para que seja aceito ( $H_0$ ) e a conclusão de não inferioridade. Na tentativa de materializar um exemplo concreto, segue abaixo a figura 1, retirada do estudo “Ensaio clínico de não inferioridade: conceitos e questões” [6].



T = tratamento; C = controle. T é superior a C se o intervalo de confiança da diferença estiver inteiramente à direita de zero, não inferior se estiver inteiramente à direita de  $-D$  e equivalente se estiver contido na zona de equivalência entre  $-D$  e  $+D$

**Figura 1** - Base teórica dos conceitos aplicados durante o ensaio clínico não-randomizado

O momento ideal para recorrer a estudos de não inferioridade é quando surge uma nova intervenção ou quando é necessário testar algo novo. Contudo, é necessário oferecer vantagem explícita em relação à intervenção já consagrada na literatura, justificando a aceitação da não inferioridade. Um exemplo de intervenção que, em nosso entendimento, justificaria o teste de não inferioridade é o treinamento intervalado de alta intensidade (HIIT) no resultado de melhoria do  $VO_{2\text{máx}}$ . Uma vez demonstrado num desenho de não inferioridade que o HIIT pode atingir o limiar de não inferioridade, podemos concluir que a intervenção é de facto eficiente em termos de tempo.

Na tabela I, foi reimpresso o estudo de Pinto [6], em 2010, no qual o autor apresenta um algoritmo de análise para três tipos de estudos hipotéticos, em que T representa a medida de eficácia da nova intervenção, e C a medida de eficácia do

grupo de controle. Rejeitar a hipótese nula significa, para estudos de superioridade, que a nova intervenção denominada T é superior ao grupo controle C. Estudos de não inferioridade, quando a diferença entre C e T é menor que uma margem delta (D) margem de não inferioridade e, para estudos de equivalência, que a diferença entre C e T não é menor nem maior que uma margem D. Fundamentalmente, o termo equivalente significa não inferior e não superior, e o teste de equivalência refere-se à análise da região simétrica definida por  $[+D,-D]$ .

**Tabela I** - Formulação de hipóteses para estudos de superioridade, não inferioridade e equivalência

Desenho do Estudo	Hipótese Nula	Hipótese alternativa
Superioridade	$H_0: C-T \geq 0$	$H_a: C-T < 0$
Não-inferioridade	$H_0: C - T \geq D$	$H_a: C-T < D$
Equivalência	$H_0:  C-T  \geq D$	$H_a:  C-T  < D$

As medidas de eficácia são apresentadas na tabela acima como; T= nova intervenção, C =controle e D como margem de não inferioridade/equivalência

Além disso, vários fatores devem ser cuidadosamente considerados no planejamento, análise e interpretação de estudos de não inferioridade para garantir a validade interna do estudo: a) escolha da margem de não inferioridade; b) o número de participantes necessários para o estudo; c) controle da sensibilidade do estudo; d) definição da população de análise. Alguns outros factores também devem ser considerados, mas estão fora do âmbito da nossa discussão neste momento.

## Conclusão

No campo da pesquisa em Educação Física, assim como em outros segmentos do setor saúde, precisamos estar atentos para tomar as melhores decisões, e saber interpretar os resultados das descobertas científicas é uma habilidade elementar, essencial para nos tornarmos melhores profissionais. Nossa intenção foi alertar a comunidade acadêmica e os profissionais de educação física sobre a conclusão de similaridade extraída dos resultados de ensaios clínicos de superioridade de eficácia sem diferença estatística. Como frequentemente observamos, muitos profissionais justificam a aplicação de uma intervenção com base nas inferências extraídas deste modelo mental, o que torna o debate relevante e necessário

Destacamos os estudos de não inferioridade como alternativa para abordar esta questão, pois acreditamos ser uma técnica metodológica mais elegante, pois estabelece parâmetros específicos para testar a similaridade ou um limite para aceitar a não inferioridade. Para quem deseja se aprofundar no tema, recomendamos dois materiais de referência que foram valiosos na elaboração desta reflexão: um desenvolvido pelo grupo CONSORT [7] e outro fornecido pela Canadian Cancer Partnership da McMaster University [8]. Trazer este tema para uma discussão mais ampla e incentivar o desenvolvimento de diretrizes e orientações específicas sobre o assunto parece ser uma necessidade emergente.

**Potencial conflito de interesse**

Nenhum conflito de interesses com potencial relevante para este artigo foi reportado.

**Fontes de financiamento**

Não houve fontes de financiamento externas para este estudo.

## Referências

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>

