

37. Significância estatística vs. significância clínica

François Noël, junho 2026

(Revisado por José Eduardo Tanus dos Santos, MRP, USP)

Em farmacologia, é muito comum comparar um grupo controle e um grupo experimental para determinar se um fármaco induziu uma resposta, utilizando-se métodos estatísticos para avaliar se essa resposta ocorreu apenas ao acaso ou não (isto é, se é “estatisticamente significativa”). Caso a distribuição dos dados seja normal (gaussiana), a análise estatística é geralmente realizada por meio de um teste t de Student ou método semelhante, gerando um valor de P utilizado para identificar um efeito (Curtis e cols., 2015).

Curiosidade: O nome do teste vem do pseudônimo “Student”, usado por William Sealy Gosset, um estatístico que trabalhava na cervejaria Guinness, no início do século XX. Ele assinou o artigo “*The probable error of a mean*” (*Biometrika* 6(1):1-25, 1908) como Student porque a empresa não permitia que funcionários publicassem resultados científicos com o seu nome.

Diferença estatisticamente significativa

Com frequência excessiva, contudo, o valor de P é reduzido a uma dicotomia, na qual os resultados são declarados “estatisticamente significativos” quando P é igual ou inferior a um ponto de corte (geralmente 0,05 na área biomédica) e “não significativos” no caso contrário, o que se tornou uma prática estatística particularmente perniciosa (Amrhein e cols., 2019; Greenland e cols., 2016). Em razão disso, recomenda-se apresentar os valores exatos de P , e não apenas indicar se estão acima ou abaixo de 0,05 (ou de outro ponto de corte), permitindo ao leitor interpretar adequadamente valores próximos ao limiar, como 0,049 e 0,051, por exemplo.

Ademais, sabe-se que o uso do ponto de corte $P = 0,05$, tradicionalmente aceito nas ciências biomédicas, é convencional, embora amplamente criticado. Autores como Colquhoun vão (muito) além ao defender a adoção de $P < 0,001$ para reduzir substancialmente o risco de falsos positivos (apud Curtis e cols., 2015).

Em vez de focar exclusivamente no valor de P , diversos autores recomendam fortemente a apresentação do tamanho do efeito acompanhado de seus respectivos intervalos de confiança (vide abaixo).

Diferença clinicamente significativa

Ainda que as análises estatísticas indiquem uma diferença, isso não implica necessariamente que ela seja clinicamente significativa, conceito cuja definição não é trivial. De fato, em contraste com os padrões relativamente bem estabelecidos (embora discutíveis, como visto acima) para decisões relativas à significância estatística, não existem diretrizes específicas para determinar qual magnitude de diferença deve ser

considerada “cl clinicamente significativa” ou “praticamente importante” (Kaul e Diamond, 2010). Assim, o limiar de relevância clínica depende do contexto: por exemplo, uma variação de 10% pode ser altamente relevante em determinados estudos de expressão gênica, ao passo que uma redução de 90% no título viral pode ainda ser irrelevante em alguns estudos de infecção viral (Curtis e cols., 2015).

Uma vez estimada a Diferença Mínima Clinicamente Importante (DMCI) para um determinado tratamento, pode-se utilizar o intervalo de confiança do risco relativo ([Razão de risco: RR](#)) para avaliar se o tratamento oferece benefício tanto estatisticamente quanto clinicamente significativo (Fig. 1).

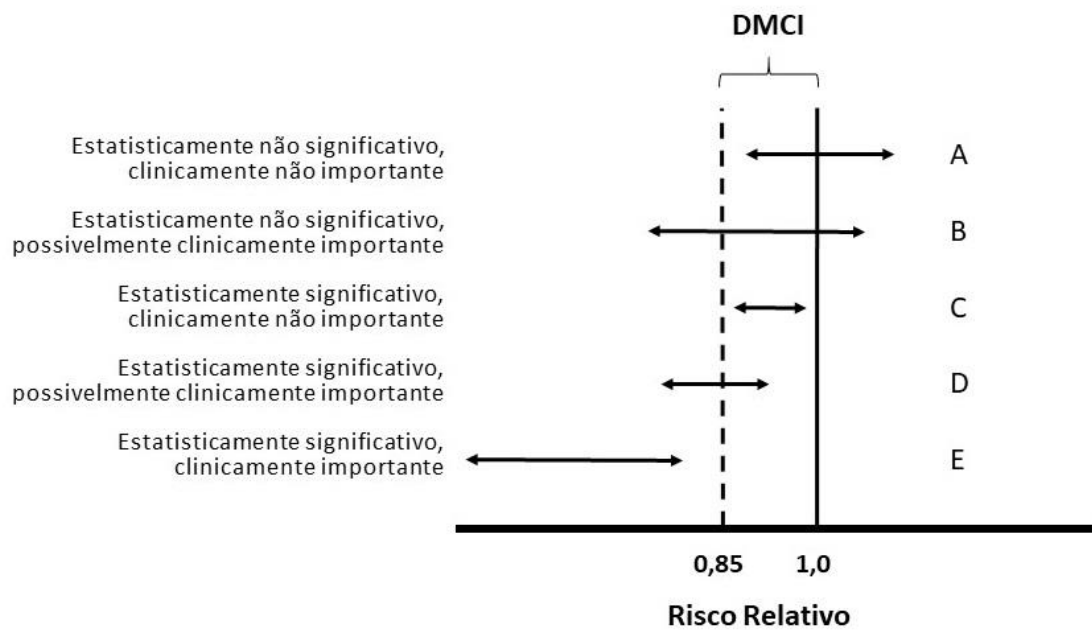


Figura 1. Representação gráfica de benefícios terapêuticos estatisticamente significativos e clinicamente importantes. São apresentados cinco resultados de ensaios (A-E) e sua interpretação em relação ao efeito nulo (razão de risco = 1,0) e à Diferença Mínima Clinicamente Importante (DMCI), correspondente a uma redução relativa de risco de 15% (equivalente a uma razão de risco de 0,85). Os efeitos do tratamento (setas duplas) são expressos como intervalos de confiança de 95%. (Adaptado de Kaul e Diamond, 2010).

Outro índice amplamente utilizado para avaliar a importância clínica de um tratamento é o [Número Necessário para Tratar](#) (NNT, equivalente a *Number Needed to Treat*). Em geral, intervenções terapêuticas são consideradas clinicamente importantes quando o NNT é < 50 (Kaul e Diamond, 2010).

Referências

Amrhein V., Greenland S., McShane B. Scientists rise up against statistical significance. *Nature* 567(7748):305-307, 2019.

Curtis M.J, Bond R.A., Spina D., Ahluwalia A., Alexander S.P.A., Giembycz M.A., e cols. Experimental design and analysis and their reporting: new guidance for publication in *BJP. Brit. J. Pharmacol.* 172:3461-3471, 2015.

Greenland S., Senn S.J., Rothman K.J., Carlin J.B., Poole C., Goodman S.N. e cols. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *Eur. J. Epidemiol.* 31:337-350, 2016.

Kaul S., Diamond G.A. Trial and error. How to avoid commonly encountered limitations of published clinical trials. *J. Am. Coll. Cardiol.* 55(5):415-27, 2010.