

## A atenção ao intervalo de confiança nas publicações



Gilson Soares Feitosa<sup>1</sup>

Um artigo recente, “The Proposal to Lower P Value Thresholds to .005”, despertou a atenção do mundo científico.<sup>1</sup>

Seu autor, o professor John P.A. Ioannidis, agora na Universidade de Stanford, levanta mais uma vez uma questão, por vezes no passado debatida por ele mesmo, enquanto atuava parte do tempo na Grécia e outra parte na Universidade de Tufts, em Boston, em 2005.

Naquele primeiro momento, ele publicou um editorial intitulado: “Why Most Published Research Findings Are False”.<sup>2</sup> Baseava-se em sua constatação de que a maioria dos trabalhos, conquanto que com apresentação de valor de  $P < 0,05$ , eram desprovidos de poder que lhes permitisse as conclusões afirmadas.

Em estudo mais recente, “Evolution of Reporting P Values in the Biomedical Literature, 1990-2015”, deteve-se em fazer detalhada pesquisa em resumos publicados em MEDLINE, além de artigos completos em PUBMED, num período de 15 anos, de 1990 a 2015. Houve um aumento de relato de valores de  $P$  de 7,3% para 15,6% em resumos e em 2014, 54,8% de ensaios clínicos randomizados demonstravam o  $P$ , invariavelmente, com achados de  $P < 0,05$  num número de aferições.

O cálculo de  $P$  diz respeito apenas à probabilidade (quando abaixo de 0,05) de negar a hipótese nula.<sup>3</sup>

Nas publicações há clara omissão de indicativos do tamanho do efeito – por exemplo, número necessário para tratar, de modo a obter um efeito, e das incertezas das medidas feitas, como a do intervalo de confiança da medida.

Agora, o autor traz à discussão o ponto de vista de que mais nos aproximaríamos da verdade dos fatos se o  $P$  fosse considerado significativo se  $< 0,005$ .

Via de regra, o caos resultante de tal medida obrigaria sempre que os estudos atuais fossem ainda mais inclusivos de pacientes, ou que vários estudos sobre o

mesmo tema devessem ser realizados para, em posterior metanálise, haver validação do achado.

Afora o fato de que veríamos, previsivelmente, de modo progressivo, de modo a justificar a nova intervenção, o ajuntamento de desfechos combinados, cada vez menos relevantes à prática clínica.


Diante do impasse surge a recomendação, até que melhor desenvolvimento apareça no futuro, de que se preste mais atenção em tamanho de efeito e medidas de imprecisão do achado, comumente avaliada pelo cálculo do intervalo de confiança que estima com 90, ou mais comumente 95%, de probabilidade, a chance de dispersão do valor do desfecho obtido.<sup>4</sup> Com tal atitude teremos uma melhor apreciação do que é visto com o valor puro de  $P$ .

Assim, quando se observa o intervalo de confiança de um resultado positivo em um ensaio deve-se levar em conta se o seu valor contém a neutralidade, ponto 0 na estimativa de comparação de uma intervenção e seu controle, e mesmo que não, pois sua distribuição é sempre maior que o ponto de neutralidade, por ser significativamente positivo, se seu menor valor possível demonstra uma grandeza que tenha relevância clínica. Caso contrário, não se deve tomar o ensaio como definitivo em si mesmo e outros serão necessários para que, se os achados forem verdadeiros, se diminua o intervalo de confiança e se obtenha maior precisão na verificação.

Por enquanto, esta parece ser a solução mais plausível para interpretação mais exigente de ensaios, ao invés de se exigir um nível de  $P < 0,005$ .

### REFERÊNCIAS

- 1- Ioannidis JPA . The Proposal to Lower P Value Thresholds to .005. JAMA 2018;319: 1429-30.
- 2- Ioannidis JPA. Why most published research findings are false. PLoS Med 2005;2(8): e124.



3- Feitosa GS. A incessante procura do “P” e seu significado em Ciência Médica. Rev. Saúde HSI 2015; 2 (set) 6-7.

4- Guyatt, G et al. Basic statistics for clinicians:2. interpreting study results: confidence intervals . Can Med Assoc J 1995; 152: 169-73.

1- Editor da Revista Científica do HSI

Endereço para correspondência:

gilson-feitosa@uol.com.br