

Medicina Baseada em Evidências – Novo Paradigma ou Pseudociência?

Marcelo Derbli Schafranski¹

DOI: 10.5935/1678-9741.20120053

RBCCV 44205-1390

É lugar-comum afirmar que o médico nunca deve parar de estudar. Concordamos em partes. De nada adianta ler e reler centenas de artigos científicos e livros-texto que se alicerçam em um modelo repleto de imperfeições, erros conceituais e armadilhas, que é a Medicina Baseada em Evidências (MBE), sem um senso crítico dos mais apurados. Previamente o que acreditávamos ser uma via segura a seguir, hoje temos como uma estrada repleta de imperfeições e que pode levar a destinos imprevisíveis, caso quem a utilize não tenha plena ciência não apenas dos erros conceituais e dos dilemas éticos aos quais a mesma está sempre exposta, mas principalmente dos artifícios metodológicos e estatísticos desse modelo, hoje quase onipresentes nos artigos científicos médicos.

Em 1998, aos nos formarmos em Medicina na Pontifícia Universidade Católica do Paraná, participávamos prazerosamente em quaisquer discussões científicas que viessem à baila, especialmente se conhecêssemos alguma meta-análise ou estudo randomizado controlado que abordasse o tema em pauta, os quais então revelávamos apenas no crepúsculo da argumentação. E o estratagema era bastante eficaz. Os mais altos níveis de evidência da emergente MBE eram raramente questionados. Os anos de experiência vieram e com eles algo que o escritor Malcolm Gladwell [1], baseado em estudos do psicólogo K. A. Ericsson [2], define como prática deliberada, indispensável para nos tornarmos ímpares dentro do que nos acostumamos a chamar de profissão. Após dez anos de atuação associados a consideráveis leituras sobre o assunto,

pudemos compreender as principais falhas e imperfeições que alicerçam a MBE como modelo para a prática e para o ensino médicos. E a observação *in loco* de constantes fracassos de terapias consideradas pela MBE como padrão-ouro nos forneceu a sustentação objetiva necessária para a hipótese que tais deficiências apresentam consequências óbvias e diretas na evolução dos nossos pacientes.

A MBE padronizou certo número de regras, difundidas em livros, apostilas e cursos espalhados pelo mundo. Mas parece que se esquece frequentemente de segui-las. A se começar pelo *P*, chamado costumeiramente de probabilidade, que recebeu a nobre missão de ter que se apresentar sempre menor ou maior do que 0,05, para que os argumentos que traz consigo possam ter ou não alguma validade científica. Dois erros conceituais nos permitem desvendar a chamada *falácia do P*. Nem mesmo seu difusor (o conceito do *P* foi desenvolvido pelo britânico Karl Pearson), o estatístico inglês R. A. Fisher, determinou enfaticamente que devesse apresentar esse valor para ter significância estatística [3]. E o pior, o valor do *P*, por menor que seja, não se refere à hipótese nula (H_0), e sim aos dados [4]. Exemplificando: imaginemos que, de maneira independente do teste estatístico que utilizamos, correto ou não, obtemos um *P* de 0,001. Conclusão (correta): uma vez que H_0 é nula, a probabilidade de nossos dados terem ocorrido é de 1 para 1000. Ou seja, H_0 será sempre falsa. E por menor que seja o *P*, Fisher jamais previu existência de hipótese alternativa (H_1). E o que, a uma primeira vista, parece apenas preciosismo conceitual, apenas aumenta ainda mais a falibilidade do *P* fisheriano: em determinadas situações, de acordo com a probabilidade pré-teste, mesmo quando o mesmo apresenta o celebrizado valor de 0,05, a chance de erro de confirmarmos uma (inexistente) H_1 pode chegar a 50%

1. Doutor em Medicina Interna. Reumatologista, Professor-adjunto de Medicina da Universidade Estadual de Ponta Grossa (UEPG) e autor de *Medicina – Fragilidades de um modelo ainda imperfeito*

[5]. Ou seja, jogar uma moeda para cima e confiar no seu resultado nos parece mais sensato, por ser igualmente “acurado” e deveras mais econômico. A confusão é tanta que chegou ao ponto do famoso intelectual Jacob Cohen indagar que se H_0 será sempre nula, qual seria a relevância em testá-la [6]. Retornamos ao problema da indução de Hume, não resolvido nem por Popper [7].

Como alternativa, os estatísticos Jerzy Neyman e Egon Pearson (filho de Karl) criaram os erros alfa (do tipo-I) e beta (do tipo-II), o primeiro não podendo ser maior do que 0,05, e o segundo, maior do que 0,2. O erro do tipo-II também é usado para calcular o poder do estudo: subtraindo-se o mesmo de 1 temos o poder estatístico da amostra, que não pode ser, por convenção, menor do que 0,8 (80%) [3]. Apesar de se tratar de abordagem dedutiva sujeita a críticas, é um modelo mais apropriado do que a questionável significância trazida pelo P . Mas a MBE transformou o erro-alfa erroneamente em P , combinando duas teorias diferentes. Ainda, para calcularmos o tamanho estimado da amostra necessária a um teste de hipóteses, usamos Neyman-Pearson, e para avaliar a validade da hipótese, o P fisheriano. Além de muitos estudos ao menos se preocuparem em calcular a amostra necessária para um adequado poder estatístico [8,9], pouquíssimos pesquisadores se lembram de calcular o poder da amostra depois de finalizada a pesquisa, o pouco conhecido *poder observado*. Resultado: estima-se que cerca de 90% dos ensaios publicados têm insuficiência de amostra, onde esses dados se encontram disponíveis [10]. Caso fossem avaliados de maneira mais criteriosa, dificilmente integrariam consensos de especialidade, e muito menos diretrizes de associações e conselhos.

Outra questão crucial envolve indicadores raramente vistos nos estudos científicos publicados: o NNT (número necessário para tratar, derivado da redução de risco absoluto – RRA, e não da redução de risco relativo, maquiagem estatística), o NNH (número necessário para causar dano) e o tamanho de efeito (*effect size*), este idealizado por Cohen, já citado aqui. São eles que nos fornecem uma pista real se o medicamento ou intervenção poderá ou não ter alguma relevância na prática médica diária. Calculadoras que os estimam abundam na *internet*, basta carregar os dados e interpretá-los. Mas não se surpreenda se encontrar intervenções cujo NNH é menor do que o NNT, ou que o tamanho de efeito se aproxima de 0 (zero), quase que igualando a terapêutica proposta ao placebo, apesar de um P significativo [11].

Mesmo que apliquemos adequadamente todos os princípios listados até aqui, ainda temos que tomar muito cuidado ao analisar o desfecho proposto pelos pesquisadores [11]. Muitas vezes, me admiro ao encontrar estudos cujo objetivo primário é algo como a redução de 30% na dor [12], ou a melhora de alguma porcentagem

insignificante no escore X, inventado pelo médico Y ou por um grupo de especialistas [13]. E todos os cálculos subsequentes se desenvolvem a partir desses desfechos pífijs. Por fim, a imponente conclusão: a terapêutica proposta é segura e eficaz (*safe and effective*). Parece que vivemos a epidemia do *safe and effective*. Digite no *Pubmed* e confira. Fosse assim, o papel da Medicina na longevidade da população seria outro.

Consideramos de natureza capital um artigo publicado em 1994, na renomada revista *Milbank Quarterly*, por um grupo de pesquisadores da *Harvard University* em conjunto com o *King's College*, de Londres, o qual deu início à nossa mudança de perspectiva em relação à verdadeira importância da evolução da Medicina durante o século XX [14]. Segundo os autores, o aumento da expectativa de vida observado ao longo do século ocorreu principalmente devido às melhorias nas condições de habitação, alimentação e saneamento básico, bem como às condições mais seguras no trânsito e no trabalho. Por meio de um método extenso e complexo, os pesquisadores concluíram que todo o avanço médico atingido durante os anos do século XX prolongaram a expectativa de vida da humanidade em algo em torno de meros cinco anos. Medidas preventivas amplamente difundidas, como triagem para hipertensão e aconselhamento para não fumar, adicionaram apenas cerca de seis meses à expectativa de vida.

Mesmo o prolongamento da expectativa de vida relacionado ao câncer é objeto de questionamento por parte de pesquisadores que estudam de maneira mais intensa as incongruências da MBE. Em sua obra *Overdiagnosis: Making people sick in the pursuit of health*, o professor Gilbert Welch [15] demonstra que a mortalidade imposta por grande parte dos cânceres, entre eles, os de mama, próstata e tireoide, se encontra estável como uma eterna linha de assistolia desde 1975, ano em que esse tipo de controle passou a ser realizado. Por outro lado, o diagnóstico dos mesmos aumenta a cada ano. Ou seja, sugere-se que estamos diagnosticando mais cedo pacientes cujo câncer não os incomodaria jamais. E pela falta de marcadores de gravidade confiáveis, os quais a Medicina ainda desconhece, inúmeros pacientes têm sido submetidos a procedimentos de risco, talvez de maneira desnecessária. As recentes polêmicas em torno da mamografia [16] e do PSA [17] se devem a esse tipo de controle estatístico, que tem sido cada vez mais divulgado até mesmo pela imprensa tida como leiga.

Seguramente a MBE tem o seu papel nessa pequena contribuição da Medicina para a longevidade da população, apesar do ufanismo da mídia, de alguns laboratórios que produzem medicamentos e equipamentos e até de determinados colegas. Cabe a nós mudarmos esse

cenário, não esperando o final de um novo século para quem sabe repetirmos os mesmos resultados obtidos com relação ao século XX. Editores e revisores de periódicos científicos não devem acolher artigos com erros e imperfeições que firam as regras da própria MBE, a qual inventou as normas e agora tem dificuldade em obedecê-las. Desse modo, embora não poderemos ter a certeza de que norteamos a evolução da Ciência Médica, certamente reduziremos o número de indivíduos expostos a tratamentos de eficácia incerta, e muitas vezes dispendiosos e perigosos.

Em 1975, o filósofo Ivan Illich abria uma de suas mais polêmicas obras com a sentença “*The medical establishment has become a major threat to health*” (“a instituição médica se tornou uma grande ameaça à saúde”) [18]. A MBE, caso seja interpretada apenas de acordo com interesses individuais ou corporativos, e não seja devidamente submetida a um escrutínio especializado, independente e, sobretudo corajoso, pode assumir o triste papel de corroboração do que foi vislumbrado pelo intelectual austríaco há mais de três décadas. Ou seja, a MBE corre o risco de se tornar mais uma dispendiosa retórica falaciosa que, além de não colaborar significativamente para a saúde da população, pode se tornar iatrogênica, pelo fato de não dispor de um controle rigoroso, acessível e confiável sobre os seus possíveis danos.

Ainda, devido à maleabilidade das ferramentas propostas pela MBE, dados idênticos podem originar evidências até mesmo antagônicas [19], cuja divulgação, parcial ou não, fica entregue ao anseio dos pesquisadores. Embora selecionar dados (“*cherry-picking*”) seja uma atitude eticamente questionável, a mesma é perfeitamente possível, em virtude do controle externo dos ensaios clínicos ainda ser incipiente [20]. Assim, a ambiguidade se incorpora ao binômio análise/resultado, colocando o modelo de conhecimento sugerido pela MBE, conforme a definição proposta pelo filósofo Karl Popper [7], no mesmo patamar de outras pseudociências, como a Astrologia.

Portanto, longe de novo paradigma, nos moldes propostos pelo filósofo estadunidense Thomas Kuhn [21], a MBE como é praticada atualmente precisa se reorganizar radicalmente para ao menos postular um lugar dentro dos limites da Ciência. E entre os desafios a serem superados, urgem a adequação de posturas éticas, o abandono de conceitos consagrados errôneos e a normatização do que deve ser realmente valorizado, publicado e difundido nos artigos científicos, conduta esta que também inclui a rigorosa subtração de análises irrelevantes e que tenham o potencial de induzir o profissional a condutas impróprias avalizadas pela MBE. Por ora, voltemos a valorizar os conhecimentos anatômico, fisiopatológico e farmacológico genuínos, assim como o bom senso e a experiência trazidas pelos anos, estes indiscutivelmente os pilares de uma

Ciência que sobreviveu por mais de dois mil anos sem ao menos um medicamento ou conduta agraciados pelas dogmáticas e recém-comoditizadas evidências.

REFERÊNCIAS

1. Gladwell M. *Outliers: the story of success*. New York:Little, Brown and Company;2008.
2. Ericsson KA, Krampe RT, Tesch-Römer C. The role of deliberate practice in the acquisition of expert performance. *Psycholog Rev*. 1993;100(3):363-406.
3. Ziliak ST, McCloskey DN. The cult of statistical significance: how the standard error costs US jobs, justice, and lives. *Ann Arbor:University of Michigan Press;2008*.
4. Vickers AJ. *What is a p-value anyway? 34 stories to help you actually understand statistics*. Boston:Addison Wesley;2009.
5. Motulski H. *Intuitive biostatistics*. Oxford:Oxford University Press; 2010.
6. Cohen J. The earth is round ($p<0.05$). *Am Psychol* 1994;49(12):997-1003.
7. Valle B, Oliveira PE. *Introdução ao pensamento de Karl Popper*. Curitiba:Champagnat;2010.
8. Leguisamo CP, Kalil RAK, Furlani AP. Efetividade de uma proposta fisioterapêutica pré-operatória para cirurgia de revascularização do miocárdio. *Rev Bras Cir Cardiovasc*. 2005;20(2):134-41.
9. Manrique R, Pavanello R, Magalhães HM. Emprego da nimodipina (oxigen) como protetor cerebral na cirurgia de revascularização do miocárdio com circulação extracorpórea em pacientes idosos. *Rev Bras Cir Cardiovasc*. 1996;11(4):248-58.
10. Lochner HV, Bhandari M, Tornetta P 3rd. Type-II error rates (beta errors) of randomized trials in orthopaedic trauma. *J Bone Joint Surg Am*. 2001;83-A(11):1650-5.
11. Gauch R. *It's great! oops, no it isn't: why clinical research can't guarantee the right medical answers*. Nova Iorque:Springer;2008.
12. Häuser W, Wolfe F, Tölle T, Uçeyler N, Sommer C. The role of antidepressants in the management of fibromyalgia syndrome: a systematic review and meta-analysis. *CNS Drugs*. 2012;26(4):297-307.

-
13. Goldman JG, Goetz CG, Brandabur M, Sanfilippo M, Stebbins GT. Effects of dopaminergic medications on psychosis and motor function in dementia with Lewy bodies. *Mov Disord.* 2008;23(15):2248-50.
 14. Welch HG, Schwartz L. *Overdiagnosed: making people sick in the pursuit of health.* Boston:Beacon Press;2011.
 15. Bunker JP, Frazier HS, Mosteller F. Improving health: measuring effects of medical care. *Milbank Q.* 1994;72(2):225-58.
 16. Welch HG, Frankel BA. Likelihood that a woman with screen-detected breast cancer has had her “life saved” by that screening. *Arch Intern Med.* 2011;171(22):2043-6.
 17. Moyer VA. Screening for prostate cancer: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med.* 2012;157:11-6.
 18. Illich I. *Limits to medicine: medical nemesis, the expropriation of health.* London: Marion Boyars Publishers;1975.
 19. Rucker G, Schumacher M. Simpson's paradox visualized: the example of the Rosiglitazone meta-analysis. *BMC Med Res Methodol.* 2008;8:34.
 20. Califf RM, Zarin DA, Kramer JM, Sherman RE, Aberle LH, Tasneem A. Characteristics of clinical trials registered in ClinicalTrials.gov, 2007-2010. *JAMA.* 2012;307(17):1838-47.
 21. Kuhn T. *The structure of scientific revolutions.* Chicago: University of Chicago Press;1996