Evidence-Based Medicine - New paradigm or pseudoscience?

Marcelo Derbli Schafranski[1]

It is commonplace to say that the doctor should never stop learning. We partly agree with it. There is no use in reading and rereading hundreds of scientific articles and textbooks which are based on a model full of imperfections, misconceptions and pitfalls, which is the Evidence-Based Medicine (EBM), without a critical sense of these information. Previously believed to be a safe route to follow, as we now have a road full of imperfections that can lead to unpredictable destinations, if used by someone who is not fully aware not only of conceptual errors and ethical dilemmas to which it is always exposed, but mainly of the methodological and statistical artifacts of this model, now almost ubiquitous in medical papers.

In 1998, we graduated from Medicine at the Pontifical Catholic University of Paraná, we would joyfully take part in any scientific discussion, especially if we knew a meta-analysis or randomized controlled trial that addressed the topic under discussion, which we would only reveal during the argument. This trick was quite effective. The highest levels of evidence of the emerging EBM were rarely questioned. The years of experience came and along with it the writer Malcolm Gladwell [1], based on studies of the psychologist K. A. Ericsson [2], defined as a deliberate practice, which is essential for us to become unique in what we are accustomed to call the profession. After ten years of operation associated with considerable reading on the subject, we could understand the major flaws and imperfections that underpin the EBM as a model for practice and for teaching physicians. And the *in loco* observation of constant failures of therapies considered by EBM as the gold standard provided us with the necessary objective support for the hypothesis that these deficiencies have obvious and direct consequences in the evolution of our patients.

EBM standardized number of rules, disseminated in books, textbooks and courses all around the world. But it seems that often forgets to follow them. Initially, the P, usually called the probability that was given the noble task to of having to always be smaller or larger than 0.05, so that the arguments it brings may or may not have any scientific validity. Two conceptual errors allow us to unravel the so-called fallacy of P. Not even its diffuser (the P concept was developed by the British Karl Pearson), the English statistician R. A. Fisher emphatically ruled that we should have this value for statistical significance [3]. It is worse than that, the value of P, however small, does not refer to the null hypothesis (H0), but the data [4]. For example: imagine that, independently on the statistical test we use, correct or not, we get a P of 0.001. Conclusion (right): once H0 is zero, the probability of our data has occurred is 1 to 1000. In other words, H0 is always false. And no matter how small the P is, Fisher never predicted the existence of alternative hypothesis (H1). At first sight, it appears only conceptual preciosity, because it only adds to the unreliability of the P proposed by Fish: in some situations, according to pre-test probability, even when it has the value of 0.05, the chance to confirm an error (nonexistent) H1 can reach 50% [5]. Tossing a coin and trust its result seems more sensible to be equally "accurate" and indeed more economical. This is so confusing that made the famous intellectual Jacob Cohen wonder why it would be relevant to test it if H0 is always zero. [6] We return to Hume's problem of induction, which could not even be solved by Popper [7].

1. Doctor of Internal Medicine at UFPR; Assistant Professor.

Alternatively, the statistician Jerzy Neyman and Pearson Pearson (Karl's children) created the alpha errors (type-I) and beta (type II), the first may not be greater than 0.05, and the second greater than 0.2. The type II error is also used to calculate the power of the study: subtracting 1, we have statistical power of the sample, which cannot be, by convention, less than 0.8 (80%) [3]. Although it is subject to criticism deductive approach, it is a more appropriate model than the questionable significance brought by P. But the EBM turned the alpha error mistakenly into P, combining two different theories. Still, to calculate the estimated sample size needed for a hypothesis test, Neyman-Pearson is used, and to assess the validity of the hypothesis, P designed by Fish. Many studies does not even bother to calculate the sample size required for adequate statistical power [8,9], and few researchers remember to calculate the power of the sample after the final survey, the little-known *observed power*. Result: It is estimated that about 90% of published trials have insufficient sample, where such data are available [10]. If they were evaluated more carefully, they would hardly integrate specialties consensus, association guidelines and councils.

Another key issue involves indicators rarely seen in published scientific studies, the NNT (number needed to treat, derived from the absolute risk reduction - ARR, and not the relative risk reduction, statistical makeup), the NNH (number needed to harm) and effect size, which was idealized by Cohen, previously mentioned here. They are the ones who give us a real clue that the drug or intervention may or may not have any relevance in daily medical practice. The calculators that are used to estimate them can be found on the Internet, all you need to do is simply load the data and interpret them. But don't be surprised if you find interventions whose NNH is smaller than the NNT or the effect size approaches 0 (zero) almost matching the proposed therapy with placebo, despite a significant P [11].

Even if we apply all the principles properly listed here, we still have to be very careful when analyzing the outcome proposed by researchers [11]. I have already found studies whose primary objective is something like 30% reduction in pain [12], or the improvement of some insignificant percentage in some scores, invented by some doctor or a group of specialists [13]. And all subsequent calculations are developed from these negligible outcomes. Finally, here comes the big conclusion: the proposed therapy is safe and effective. It seems that we live in the safe and effective epidemic. Type it in on Pubmed and check it out. If that were true, the role of medicine in the longevity of the population would be totally different.

The article published in 1994 in the renowned journal Milbank Quarterly, by a group of researchers from Harvard University in conjunction with King's College from London, began our change of perspective in relation to the true importance of the evolution in medicine during the twentieth century [14]. According to the authors, the increase in life expectancy observed throughout the century was mainly due to improvements in housing conditions, nutrition and sanitation, as well as safer conditions in traffic and at work. Through an extensive and complex method, the researchers concluded that the entire medical breakthrough achieved during the years of the twentieth century extended the human life span into mere five years. Widespread preventive measures such as screening for hypertension and advice not to smoke, added only about six months to life expectancy.

Even the extension of life expectancy related to cancer is a subject of inquiry by researchers studying more intensively the inconsistencies of EBM. In his book "Overdiagnosis: Making people sick in the pursuit of health", the Professor Gilbert Welch [15] shows that the mortality imposed by the majority of cancers, including the breast, prostate and thyroid cancers, is stable as an eternal asystole line since 1975, year that this type of control started to be performed. On the other hand, their diagnosis increases every year. We may affirm that, we have been diagnosing early patients whose cancer would never bother them. The lack of reliable markers of severity, which is still unknown by Medicine, many patients have undergone risk procedures, perhaps unnecessarily so. Recent controversies surrounding mammography [16] and PSA [17] are due to this type of statistical control, which has been increasingly reported by the press that is considered not to have any knowledge about the subject.

The EBM has surely its role in this small contribution of Medicine for the longevity of the population, despite the media boosterism of some laboratories that produce drugs and equipment and even some colleagues. It is up to us to change this scenario, not expecting the end of a new century to perhaps repeat the same results with respect to the twentieth century. Editors and reviewers of scientific journals should not accept articles with errors and flaws that contradict the rules of EBM itself, which invented the rules and now has difficulties in following them. Thus, although we cannot be sure that guiding the evolution of medical science certainly will reduce the number of individuals exposed to treatments of uncertain efficacy, and often costly and dangerous.

In 1975, the philosopher Ivan Illich opened one of his most controversial works with the sentence "The medical establishment has become a major threat to health" [18]. The EBM, if interpreted only according to individual or corporate interests, and not properly subject to a specialized scrutiny, independent and above all courageous, can take the sad role of corroboration of what was envisioned by the Austrian intellectual for more than three decades. In

other words, EBM is likely to become a fallacious rhetoric that, besides not significantly contributing to the population health, can become iatrogenic, because they do not have a strict control, accessible and reliable information on their possible damage.

Still, due to the flexibility of the tools proposed by EBM, the identic data can even create antagonistic evidences [19], and its partly and/or complete publication is delivered to the desire of researchers. Although selecting data ("cherry-picking") is an ethically questionable attitude, it is perfectly possible, due to the incipient external control of clinical trials [20]. Thus, the ambiguity is incorporated into the binomial analysis / results, putting the knowledge modeling suggested by MBE, as the definition proposed by the philosopher Karl Popper [7], at the same level as other pseudo-sciences as astrology.

Therefore, far from a new paradigm, as proposed by the American philosopher Thomas Kuhn [21], the EBM, as it is currently practiced, needs to be radically reorganized to, at least, be respected within the limits of science. Among the challenges to be overcome, the appropriateness of ethical stances are urged, the abandonment of erroneous concepts and normalization of what should be really valued, published and distributed in scientific articles, and the rigorous subtraction of irrelevant analysis that may induce professional misconduct endorsed by EBM. We should now value the anatomical, pathophysiological and pharmacological genuine knowledge, as well as common sense and experience brought by the years, which are the pillars of a science that has survived for more than two thousand years without at least one medication or conduct graced by dogmatic and commoditized recent evidences.

REFERENCES

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