
Application of the Principles of Evidence-Based Medicine in Programs to Assess High-Cost Procedures and Medications

Edimar Alcides Bocchi, José Antonio Marin Neto

São Paulo, Ribeirão Preto, SP - Brazil

Determining the efficacy of treatments and procedures is of fundamental relevance, because economic resources are limited and, therefore, social priorities should be established¹. Despite the benefit resulting from the continuous development of new drugs that reach clinical success, this process is frequently accompanied by elevation in costs due to the legitimate need for remuneration and return of investments in the area of pharmaceutical research.

Physicians and private and public institutions responsible for providing health programs often have to face the dilemma of replacing or not medications already in use. Likewise, the dilemma of adding new drugs or not to traditional therapeutical protocols is frequently faced, with no spurious influence of interests void of scientific and ethical basis, aiming at obtaining a socially adequate cost/benefit ratio.

The process by which changes in medical practice are introduced and authorized has undergone transformations in its fundamentals². Two major innovations, the randomized clinical trials and the guidelines provided by specialized medical societies, shifted the old emphasis on clinical application of the experience in individual practice to the use of scientific evidence-based norms. These essential tools that guide the best currently individualized medical practice should also be used to optimize the application of resources of the community in the issue in question, to guide the adoption of high-cost medications and procedures³.

The practice of evidence-based medicine consists of harmonization and integration of the better external clinical evidence obtained through systematic research about studies specifically directed to obtaining conclusive results, which are usually gathered as guidelines/consensus, with

the personal experience and ability to individualize treatments and medical practices for the patients who require them. These tools are synergistically complementary. Great clinical trials with randomization of treatment for groups of patients provide better evidence, which may be systematically revised and critically considered through the guidelines. This is the arena where the practical importance of the studies may be finally assessed and the extension of their employment in the population of patients may be determined. Both should be impersonal, objective, and void of nonscientific interests, and these are crucial characteristics of their power to convince. Individual medical ability derives from personal experience, from the capacity for clinical judgment, from a critical observation of the medical literature, and from the participation and discussion in medical meetings. When specialists with these qualities are gathered under the sponsorship of governmental entities or interested medical societies, or both, a conjunct position is determined, more or less consensual, in guidelines that reflect the set of these tools of modern normativeness of medical management⁴. Usually, each procedure is classified in classes according to the level of evidence supporting it (Table I).

Elaboration of guidelines for medical practice may reduce variability in medical practice and, consequently, optimize the cost/benefit ratio⁵. However, the isolate application of these guidelines may be subjected to individual and social limitations, among others, and should be integrated with other tools for more impacting decisions. The use of evidence obtained in clinical studies developed in a systematic manner should constitute the major support for changing effective decisions that may be widely applied to medical literature. The possibility of recommending therapeutical procedures always requires a degree of consistency of the evidence obtained from studies, and the procedures can be graded according to the methodology used (Table II)^{6,7}. In the absence of large or multiple randomized trials, the best evidence available is searched and hierarchized according to the principles exposed. Small studies, even though randomized, are usually useful to elucidate mechanisms, and

Instituto do Coração do Hospital das Clínicas da FMUSP and Faculdade de Medicina de Ribeirão Preto da USP

Mailing address: Edimar Alcides Bocchi - Rua Oscar Freire, 2077/161 - 05409-011 - São Paulo, SP, Brazil

English version by Stela Maris C. Gandour

Table I – Classification of the procedures

Class I	
Conditions for which conclusive evidence exists, or in its absence, general consensus that the procedure is useful or effective, or both.	
Class II	
Conditions for which conflicting evidence or divergence of opinion, or both, exists in regard to usefulness/efficacy of the procedure.	
Class II A	
Weight or evidence/opinion favoring usefulness/efficacy.	
Class II B	
Usefulness/efficacy less well-established by evidence or opinion.	
Class III	
Conditions for which evidence or consensus, or both, exists that the procedure is not useful/efficient, and, in some cases, it may even be noxious.	
Adapted from the criteria used in the guidelines of the American College of Cardiology/American Heart Association.	

Table II – Grading of recommendations and levels of evidence

Grade A	
Level	1a) Evidence of randomized clinical trials including a large number of patients, or a systematic review including meta-analysis of several medium-sized randomized studies, with a total number of patients comparable to that of the first situation.
Level	1b) Evidence of at least 1 high-quality “all/nothing” population study in which all patients with the conventional therapy die or are unsuccessful (nothing) and a few/several patients survive or are successful with the new therapy (all) (ex.: chemotherapy for tuberculosis, meningitis, defibrillation for ventricular fibrillation, etc), or studies in which the conventional therapy fails or is associated with death (all) and no patient dies or is unsuccessful with the new therapy (nothing) (ex: penicillin for pneumococcal infection, antibiotic therapy for infectious endocarditis, etc).
Level	1c) Evidence of at least 1 medium-sized clinical trial or meta-analysis of studies with a small number of patients, which together sum up a moderate number of patients.
Level	1d) Evidence of at least 1 randomized clinical trial.
Grade B	
Level	2) Evidence of at least 1 high-quality study with a nonrandomized population, to receive or not the new therapeutics.
Level	3) Evidence of at least 1 high-quality study with control group.
Level	4) Evidence of at least 1 high-quality study involving a series of patients with no control group.
Grade C	
Level	5) Opinions of specialists with no reference, or use of already referred studies, or of physiopathological or experimental bases. A comprehensive evaluation should consider multiple types of evidence (ex.: randomized studies, nonrandomized, epidemiological and experimental observations) and check the framework of information for consistency, coherence, clearness, and specific subgroups of use. In some circumstances, results of nonrandomized studies are so clear and biologically plausible that they could reasonably be considered grade A studies (ex.: artificial pacemaker for total atrioventricular block).
Adapted from Yusuf S, Cairns JA, Camm JA, Fallen EL and Gersh BJ, editors, Evidence Based Cardiology. London: Br Med J Books, 1998.	

generate and support hypotheses, mainly when grouped in consistent meta-analyses. The essential issues raised in this way are finally scrutinized by randomized studies with a large number of patients, allowing the exploration of incidence, even of those with less frequent outcomes but in-

fluencing clinical practice (ie, mortality), in order to eliminate contamination of factors spurious to treatment, allowing totally reliable results^{8,9}. The consequences of new treatments should be measured by results that support or not, in an objective and conclusive manner, their application and repercussion in the cost/benefit ratio, which is significant in high-cost methods of treatment.

Comparison of the consequences of the introduction of new drugs and identification of the patients who are at high risk of an event and may respond to these new forms of treatment is paramount. The most used indices for measuring efficacy or clinical benefit are as follows: relative reduction in risk, odds ratio, absolute reduction in risk, and number of patients to be treated^{10,11} (Table III). The major limitation of the use of the relative reduction in risk (most common because it is expressed in a very impressive numerical form) is that it does not reflect the absolute magnitude of the risk without treatment, and it may overestimate or underestimate the absolute impact of the treatment when events are very rare or very frequent in the groups. The odds ratio expresses calculated risks in retrospective studies (usually of the case-control type) and, when events are rare, it may be an adequate estimate of the relative risk, which is immediately extracted from the results of prospective studies. The number of patients required for a certain treatment expresses the number of patients who should be treated to avoid an event as compared with that of patients who are not treated in the same way of the study. If the rate of event is high, even with a relatively small reduction in the risk, a small number of patients will be large enough for obtaining the benefit of avoiding a clinically relevant event.

Clinical trials with randomized allocation of patients for each modality of treatment are powerful tools for testing new therapeutics, but several points require consideration in the analysis of results^{12,13}. Fundamental limitations exist, as do limitations of study applicability. In regard to fundamental limitations, the following stand out: 1) restriction of the period of treatment and limitation of the follow-up period: when the follow-up period is short, even though an initial beneficial result may exist, later this effect may be lost. The period of treatment is important. Extrapolation of the results may not be correct, if performed routinely¹⁴; 2) absence of tests with different drug dosages and combinations of medications

Table III - Measurements of clinical benefit

1) Relative reduction in risk (RRR)	$\frac{\text{rate of events in the control group} - \text{rate of events in the group treated}}{\text{rate of events in the control group}}$
2) Odds ratio (OR)	$\text{Chance ("Odd")} = \frac{P(\text{probability}) \text{ of an event}}{(1-P)}$
Odds ratio (OR) =	$\frac{\text{Chance in the control group}}{\text{Chance in the group treated}}$
3) Absolute reduction in risk (ARR) =	rate of events in the control group – rate of events in the group treated.
4) Number of patients that should be treated to avoid an event (NNT) =	reciprocal of the absolute reduction in risk = (100/ARR).

not studied; 3) validity of the results not entirely assured. The studies should be properly randomized and the significance level properly fixed in the clinical context. A result with $p < 0.05$ has a 50% chance of being replicated in another study; with $p < 0.01$, a 70% chance; and with $p < 0.001$, a 90% chance of being confirmed in a similar study. The confidence interval should be valued because of the precision it provides in estimating statistical results. Adherence to the analysis of primary objective, preferably encompassing or exclusively considering absolute and objective outcomes, such as mortality, is essential. When the main objective is improvement of morbidity, the results should be more carefully assessed; 4) size of the sample is important, from conception to the end of the randomized study. Analysis in subgroups, mainly if not previously established as a goal, and analysis of deriving studies, among others, may increase chances of error and should always be viewed with reservation; 5) replication of the results through several independent studies is also very important; 6) clinical relevance: the size of the sample should be large enough and the value of reduction in the relative risk should be assessed considering the risk of the event, especially if the procedure is expensive. If the procedure is inexpensive, this influence may be modulated and smaller reductions in risk may be acceptable; 7) low-risk outcomes: when the event rate is inherently reduced, the study power decreases and a large number of patients should be studied. Therefore, prevalence of the event in the population is decisive for planning studies with characteristics like these discussed here.

Another fundamental category of limitations of studies consists of the low applicability in medical practice, which is often neglected. Among these, we can cite: 1) the inclusion criteria of many studies drastically reduce their applicability to many patients. The possibility of final application of the results of the study depends, therefore, on the characteristics of the patients included and the treatment used in the study. In other words, a certain treatment should be applied to the appropriate patient for an adequate period of time, and in the phase correctly established in the studies. The benefit of the procedure is proportional to the risk caused by the clinical condition, and the appropriate stratification of the risk is always mandatory for the correct identification of patients benefited by the procedure. Extrapolation to medical practice is justified when the patients have the same profile as the patients included in the study, especially in regard to risk, age, sex, etiology, and severity of the clinical condition. The populations studied should be homogeneous, and only those with a real potential for benefit should be treated; 2) the results of the studies express probability of risk and benefits in means applicable to population

groups. Therefore, the reduction in relative risk would be applicable to categories of patients included in the study and the clinical decision should always be evaluated in regard to the individual characteristics of each patient; 3) treatment risks and costs should be carefully assessed in regard to potential benefits; 4) assessment of the consistency between results and biological and experimental studies; 5) the studies may, according to their own criteria of selection, not be able to solve the “gray” areas of clinical practice when the cost/benefit ratio of the options is incompletely determined or even leads to contradictory results; 6) circumstances exist where appropriate studies are not amenable to being conducted because of the involvement of unethical points or an unacceptable cost, or both; 7) most clinical trials are funded by entities (ie, pharmaceutical industry, equipment industry, etc), which have legitimate interests in the applicability of the results and, therefore, should be kept away from the processes of analysis and final publication of scientific studies. As this important principle is not always observed, studies with external financing more commonly obtain positive results, as compared with those that are not under this economical influence^{15,16}. As the interests of the manufacturers of the medication are different from those of the people who prescribe or consume it, after the finalization and publication of the scientific study, monitoring of how the product is going to be marketed is mandatory.

The guidelines or consensus may also have limitations that should be considered. The results of randomized clinical trials guide the decisions of the consensus, which guide medical practice. The consensus is a social and scientific process, in which the following points should be emphasized: 1) objective and impersonal communications; 2) all clinical evidence considered; 3) simple conclusions and direct product of the evidence; 4) even the best limited evidence and conclusions requiring judgment and interpretation; 5) some participants tend to advocate particular points of view, and this tendency should be judiciously attenuated.

Recommendations

Clinical trials and guidelines are irreplaceable and valuable. Despite their inherent limitations already discussed, they are the best tools for guiding the implementation of new therapeutical procedures. The cost/benefit ratio for social application should also be determined. To optimize the application of the limited resources of the community, evidence-based medicine should always be used in programs for implementing new and old procedures and high-cost medications.

References

1. Fletcher A, Spiegelhalter D, Staessen J, Thijs L, Bulpitt C. Implications for trials in progress of publication of positive results. *Lancet* 1993; 342: 653-7.
2. Sniderman AD. Clinical trials, consensus conferences, and clinical practice. *Lancet* 1999; 327-30.
3. Feinstein AR, Horwitz RI. Problems in the "Evidence" of "Evidence-based Medicine". *Am J Med* 1997; 103: 529-35.
4. Bocchi EA. I Diretrizes da Sociedade Brasileira de Cardiologia para Transplante Cardíaco. *Arq Bras Cardiol* 1999; 73: 5.
5. Heffner JE. Does evidence-based medicine help the development of clinical practice guidelines? *Chest* 1998; 113: 172S-8S.
6. Bjorn M, Brendstrup C, Karlsen S, Carlsen JE. Consecutive screening and enrollment in clinical trials: the way to representative patient samples. *J Cardiac Failure* 1998; 4: 225-30.
7. Kitching A, Sackett D, Yusuf S. Approaches to evaluating evidence. In: Yusuf S, Cairns JA, Camm AJ, Fallen EL, Gersh BJ. *Evidence Based Medicine*. Ed. London: Br Med J Books, 1998: 3-23.
8. Kennedy HL. The importance of randomized clinical trials and evidence-based medicine: a clinician's perspective. *Clin Cardiol* 1999; 22: 6-12.
9. Yusuf S, Kitching AD. From journal to bedside: application of clinical trial results to individual patients. *Evidence-Based Cardiovascular Medicine*, 1998; June: 29-31.
10. Laupacis A, Sackett DL, Roberts RS. Na assessment of clinically useful measures of the consequence of treatment. *N Engl J Med* 1988; 318: 1728-33.
11. McQuay HJ, Moore RA. Using numerical results from systematic reviews in clinical practice. *Ann Intern Med* 1997; 126: 712-20.
12. 1. Conti CR. Clinical Trials: Are really useful for clinical decision making? *Clin Cardiol* 1996; 19: 763-4.
13. Avezum A, Rossi-Neto JM, Piegas L. Por que precisamos de estudos randomizados e de estudos epidemiológicos em doença cardiovascular? *Cardiologia baseada em evidências VII*. *Arq Bras Cardiol* 1999; 72: 281-8.
14. Pocock S, White IAN. Trials stopped early: too good to be true? *Lancet* 1999; 353: 943-4.
15. Davidson RA. Source of funding and outcome of clinical trials. *J Gen Intern Med* 1986; 1: 155-8.
16. Stelfox HT, Chua G, O'Rourke K, Destsky DA. Conflicts of interest in the debate over calcium antagonists. *N Engl J Med* 1998; 338: 101-06.