Randomized controlled trials as sources of information in evidence-based medicine.

Seed Amir Mohsen Rezaiee MD*

Assistant professor of urology, Member of Committee
for Teaching Universities' Faculty Members Educational Development Center.

ABSTRACT

Physicians are constantly challenged by their patients, the media, and the rapid pace of technological advance to keep abreast of new diagnostic and therapeutic developments. Evidence-based medicine (EBM) is the practice of making updated medical decisions through the judicious identification, evaluation, and application of the most relevant published information. RCTs are one of the best primary sources used in EBM. Thus, evaluation of the strength and validity of the RCTs that supports the discussion on making medical decisions seems to be necessary. In this paper, main types of biases occurred during the course of an RCT as well as Assessment of RCTs as sources of information in evidence-based medicine are discussed.

Key Words: *RANDOMIZED CONTROLLED TRIAL, EVIDENCE BASED MEDICINE, ASSESSMENT*

Journal of Medical Education winter *2003;2(2):93-98*

Introduction

Physicians are constantly challenged by their patients, the media, and the rapid pace of technological advance to keep abreast of new diagnostic and therapeutic developments. The constraints of the current medical environment, with the constant need to develop cost-effective options for diagnosis and treatment, create additional challenges and leave little time for keeping up with the latest developments in medicine. It is no wonder that many physicians wish they had a deeper understanding of the benefits and risks of advances in medical technology and a better ability to determine which patients would benefit most from currently available treatments.

An estimated 4000 medical journals are published monthly worldwide, as well as 200 secondary journals that summarize the articles published in the primary journals. Physicians need special skills for gathering enough accurate information in their scant reading time to maintain the highest standards of quality in their medical practice. Here we present a systematic approach for your use in efficiently processing patient-related medical information and facilitating your decision-making capabilities.

In 1990, Gordon Guyatt, MD, currently professor on the faculty of the Health Sciences Center of Clinical Epidemiology and Biostatistics at McMaster University (Hamilton, Ontario), coined the phrase evidence-based medicine (EBM). Although, The concept of evidence-based medicine is not new. Its use has been under investigation since the mid-1980s Under Title IX of the U.S (1).

Evidence-based medicine (EBM) is the practice of making medical decisions through the judicious identification, evaluation, and application of the most relevant published information. Although the concept of practicing medicine based on sound evidence is not new, recent refinements have enabled practitioners to approach medical problems and evaluate medical literature with greater consistency and to deal with massive amounts of medical information via a qualitative approach. EBM suggests the steps needed to identify medical literature that is relevant to your practice and also helps in differentiating high quality from low quality findings (2).

In theory EBM implies that physicians make patient care decisions based on integrating individual clinical expertise with the best available external clinical evidence derived from a systematic search of medical literature The practice of EBM by physicians involves several steps

The first is to formulate an appropriate clinical question; the physician then conducts a literature search to find lab test clues that will help in finding a solution. The physician embarks on a search of the medical literature, critically appraises the to the articles found and applies the results of the search

patient's care (3).

RCTs are one of the best primary sources used in EBM. RCTs are strong evidence that exist to support the conventional approach to managing a given clinical situation. Users need to know the strength of the evidence supporting the key clinical recommendations on diagnosis and treatment. Thus, evaluation of the strength and validity of the literature that supports the discussion seems to be necessary.

It is also needed to rate only those statements that have corresponding references and base the rating on the quality and level of evidence presented in the supporting citations (4,5).

In this paper, main types of biases occurred during the course of an RCT as well as Assessment of RCTs as sources of the information in evidence-based medicine are discussed.

What is a RCT?

- An RCT seeks to measure and compare the outcomes of two or more clinical interventions.
- One intervention is regarded as the standard of comparison or control.
- Participants receive the interventions in random order to ensure similarity of characteristics at the start of the comparison.
- Randomization can be achieved through a variety of procedures.
- RCTs cannot answer all clinical questions

Typically, RCTs seek to measure and compare different events that are present or absent after the participants receive the interventions. These events are called outcomes.

In summary, RCTs are quantitative, comparative, controlled experiments in which a group of investigators studies two or more interventions in a series of individuals who receive them in random order (6).

Randomized controlled trials (RCTs), even if assembled into a perfect systematic review, are just one of the many different types of information that can inform decisions.

A perfect RCT

To be perfect, trials would have to do the following:

- Answer clear and relevant clinical questions previously unanswered.
- Evaluate all possible interventions for all possible variations of the conditions of interest, in all possible types of patients, in all settings, using all relevant outcome measures.
- Include all available patients.
- Include strategies to eliminate bias during the administration of the interventions, the evaluation of the outcomes, and reporting of the results, thus reflecting the true effect of the interventions.
- Include perfect statistical analyses.
- Be described in reports written in clear and unambiguous language, including an exact account of all the events that occurred during the design and course of the trial, as well as individual patient data, and an accurate description of the patients who were included, excluded, withdrawn, and dropped out.
- Be designed, conducted, and reported by researchers who did not have conflicts of interest.
- Follow strict ethical principles (7).

If all trials were perfect, you would not have to worry about their quality. Instead, you could always use them with confidence as part of your decisions. Unfortunately, there is no such as thing as a perfect trial. In real life, readers only have imperfect trials to read and face lots of barriers to determining their quality with confidence.

What is quality in evaluating of RCTs?

Quality means different things to different people. For example, clinicians used to evaluate a RCT quality according its external validity and usually seek for the generalisability of the results because they trying to decide whether to offer its result to another patients. Researchers are supposed to focus on internal validity of RCTs as an indicator for the quality because they almost look at RCTs as references for the future studies (4,5).

a peer-reviewer may evaluate RCTs quality by sing other aspects of the trials because he trying to decide whether to recommend the report for publication.

Specific aspects of trials that have been used to define and assess trial quality include the following:

- The clinical relevance of the research question.
- The internal validity of the trial
- The external validity (the precision and extent to which it is possible to generalize the results of the trial to other settings).
- The appropriateness of data analysis and presentation.
- The ethical implications of the intervention they evaluate.

Of all the aspects of a trial that have been used to

define and assess quality, internal validity is the least context-dependent and perhaps the only one that has been the subject of the few empirical methodological studies available. As a result of this, it is recommend that elements related to internal validity should be included in any assessment of the trial quality.

The internal validity of the trial is defined as the degree to which the trial design, conduct, analysis, and presentation have minimized or avoided biased comparisons of the interventions under evaluation. it is therefore, any trial that have an acceptable internal validity can produce valid results(7).

Any factor that influence measurements and evaluations randomly is named random error but bias is defined as any factor or process that tends to deviate the results or conclusions of a trial systematically away from the truth.

Influences of random errors can be reduced by increasing sample size of the trial but it can not be avoided completely, but discussions on bias should focus on all biases that can occur at any point during the course of a trial and , any empirical methodological studies that support their existence and the methods that are have been implied to control the biases in different trials .

there are many biases that relate to RCTs (table 1) but in this paper, only the main biases that can occur in RCTs are discussed(8).

Table 1 BIASES THAT CAN OCCUR IN RCTS

Main types of bias in RCTs
Selection bias
Ascertainment bias
Other biases:
Bias introduced by inappropriate
Handling of withdrawals, drop out
Publication bias
Language bias
Country of publication bias
Time lag bias
Potential breakthrough bias
Biases can occur during the
Uptake of information by users:
Rivalry bias
I owe him one bias
Personal habit bias
Moral bias
Clinical practice bias
Territory bias
Complementary medicine bias
Do something bias
Do nothing bias
Favored design bias
Disfavored design bias
Resource allocation bias
Prestigious journal bias
Non-prestigious bias
Printed word bias
Prominent author bias
Unknown or non-prominent author bias
Famous institution bias
Unrecognized or non-prestigious institution bias
Large trial bias
Multicenter trial bias
Small trial bias
Flashy title bias
Credential or professional bias
Esteemed author bias
Geography bias
Tradition bias
Bankbook bias
Belligerence bias
Technology bias
Empiricism bias
I am an epidemiology bias

Selection bias

In a true design, all the study participants are given the same opportunity to be allocated or assigned to each of the study groups. Selection bias occurs when the outcomes of a trial are affected by systematic differences in the way in which individuals are accepted or rejected for a trial or in the way in which the interventions are assigned to individuals once they have been accepted into a trial.

For example, in a trail performed for comparing open prostatectomy and Transurethal Resection of the Prostate (TURP), if all high risk patients are assigned to the group treated by TURP or all patients with large cysts are assigned to the group treated by open prostatectomy.

It is obvious that this trial will not only show the effects of surgical intervention. The finding of the trail are the result of the surgical intervention and allocation of participants both. Therefore, such trail can not be used by a clinician.

The best way to protect a trial against selection bias is by randomization of Individuals, groups, and the order in which measurements are obtained.

Ascertainment bias

Ascertainment bias occurs when the results or conclusions of a trial are systematically distorted by knowledge of which intervention each participant is receiving. Ascertainment bias can be introduced by the person administering the

interventions, the person receiving the interventions (the participants), the investigator assessing or analyzing the outcomes, and even by the people who write the report describing the trial. For instance, in a trail about therapeutic effects of a new drug , the new drug may appear to be more effective at the end of the trial

if participants know that they have been allocated to the placebo group, they are likely to feel disappointed and less willing to report improvement at each of the study time-points. In addition, if the people in charge of assessing and recording the outcomes know which patients are allocated to each of the study groups, they could, consciously or unconsciously, tend to record the outcomes for patients receiving the new drug in a more favorable way than for patients receiving placebo.

The best way to protect a trial against ascertainment bias is by keeping the people involved in the trial unaware of the identity of the interventions for as long as possible. This is also called blinding or masking.

The strategies that can be used to reduce ascertainment bias can be applied.

The importance of blinding has been confirmed in empirical studies. It has been shown, for instance, that open studies are more likely to favor experimental interventions over the controls and that studies that are not double-blinded can exaggerate effect estimates by 17%(9). Despite the empirical evidence available, and common sense, it has been shown recently that only about half of the trials that could be double-blinded actually achieved double-blinding(10) Even when the trials are described as double-blind, most reports do not provide adequate information on how blinding was achieved or statements on the perceived success (or failure) of double-blinding efforts.

The best strategy to achieve blinding during data collection is using of placebos.

Biases caused by participants' withdrawal or inappropriate Handling of participants'withdrawals and drop out .

Ideally, all participants in a trial should complete the study, follow the protocol, and provide data on all the outcomes of interest at all time-points. In reality, however, most trials have missing data. Data can be missing because some of the participants drop out before the end of the trial, because participants do not follow the protocol either deliberately or accidentally, or because some outcomes are not measured correctly or cannot be measured at all at one or more time-points. Regardless of the cause, inappropriate handling of the missing information can lead to bias. For instance, if in the new drug trial patients who do not obtain benefit from the new drug withdraw more frequently because of adverse effects, their exclusion from analysis would lead the investigators to exaggerate the benefit and underestimate the harm of the new drug.

Sometimes participants move to different areas during the study or fail to contact the investigators for an unknown reason. Excluding these participants or specific outcome measurements from the final analysis can also lead to bias.

The only strategy that can confidently be assumed to eliminate bias in these circumstances includes two components. The first is called 'intention to treat' analysis, and means that all the study participants are included in the analyses as part of the groups to which they were randomized regardless of whether they completed the study or not. For example, if a participant in TURP group was high risk for heart diseases, he/she should not be allocated to open prostatectomy group.

The second component includes a 'worst case scenario' sensitivity analysis. This is performed by assigning the worst possible outcomes to the missing patients or time-points in the group that shows the best results, and the best possible outcomes to the missing patients or time-points in the group with the worst results, and evaluating whether the new analysis contradicts or supports the results of the initial analysis which does not take into account the missing data(8).

Assessment of RCTs' quality

There are many tools and methods for assessment of RCTs quality. One of these tools has been developed by Jaded .it is easy and quick to use and its Validity have been well documented previously (11,12,13,14).in this tools, assessments are based on existence of 3 main biases pointed in this paper in RCTs and methods implied for controlling these biases.

This tool is a scale that includes three items that are directly related to bias reduction and are presented as questions to elicit 'yes' or 'no' answers. The scale produces scores from 0 to 5. Point awards for the first two items (randomization and double-blinding) depend not only on whether the trial is described as randomized or double-blind, but also on the appropriateness of the methods used to randomize and blind the trial. For example, if the trial is described as randomized or double-blind, but there is no description of the methods used to generate the randomization sequence or the double-blind conditions, 1 point is awarded in each case (that is, 1 point is awarded

for randomization and 1 point for double-blinding if the trial is described as both randomized and double-blind, or only 1 point is awarded if the trial is described as randomized but not as double-blind). If the methods of generating the randomization sequence or creating blinded conditions are described and are appropriate, 1 additional point is given for each item. Conversely, if the methods used to generate the randomization sequence or create blinded conditions are described, but inappropriate, the relevant item is given 0 points. The third item of the scale, withdrawals and dropouts, is awarded 0 points for a negative answer and 1 point for a positive answer. For a positive answer, the number of withdrawals and dropouts in each group and the reasons must be stated in the report. If there were no withdrawals, this should also be stated (Fig. 1). If a trial provides the number and reasons for withdrawals and dropouts in each group, you, as a reader, could reanalyze the data. At the time of the development of the scale, it was debated whether this item should be scored according to the proportion of withdrawals and drop outs in the trials, but this was considered inappropriate because we do not know precisely when a trial has too many drop outs. Once you have scored all the items of the scale, a trial could be judged as having poor quality if it is awarded 2 points or less.

A checklist is also presented in this paper to bring a scale for assessing of other aspects of a RCT to the attention of the readers (table 2).

Perhaps assessment of RCTs quality is difficult to do especially for those who have not adequate knowledge in the field of RCTs quality assessment. there are several comprehensive source of RCTs that present RCTs with acceptable quality.

The Cochrane Library is the most advanced and comprehensive source of evidence, especially from RCTs, needed to make informed health care decisions.

the Cochrane Controlled Trials Database contains citations for thousends of controlled trials identified through the collective effort of members of the Cochrane Collaboration to improve the identification of primary studies. More information on this database can be obtained from the Internet at:http://hiru.mcmaster.ca/cochrane.

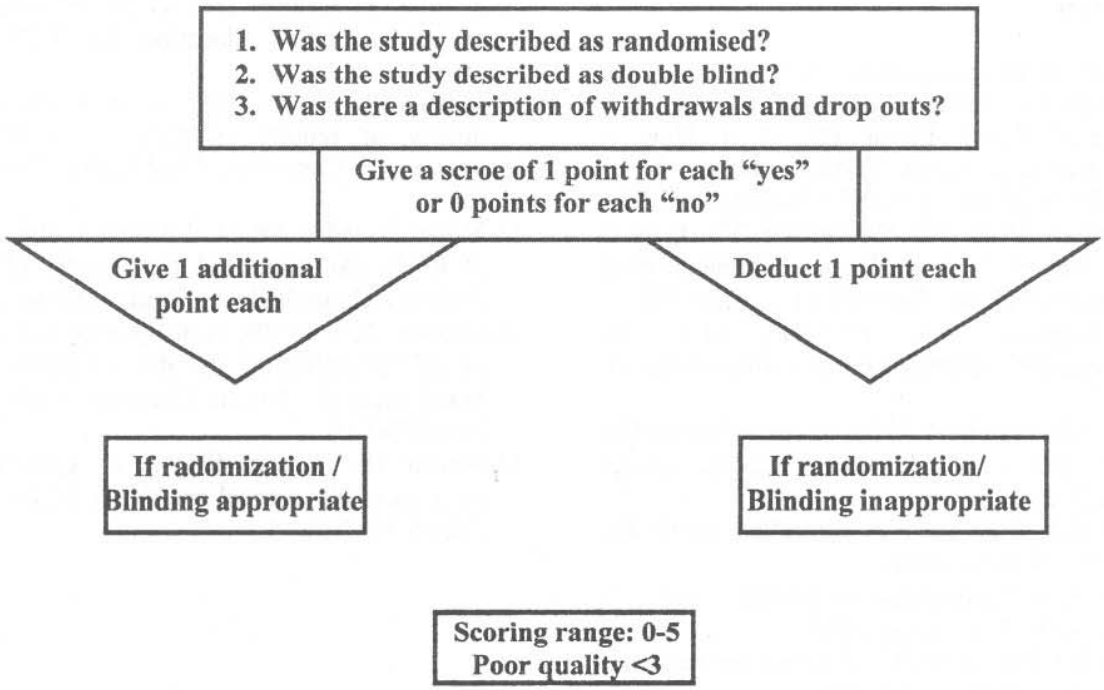

Fig 1 VALIDATED QUALITY SCALE. (FROM JADAD ET AL.)

Table 2 A CHECKLIST FOR APPRAISING RANDOMIZED CONTROLLED TRIALS

1-was the objective of the trial sufficiently described?
2-was a satisfactory statement given of the diagnostic criteria for entry to the trial?
3-were concurrent controls used (as opposed to historical controls)?
4-were the treatments well defined?
5-was random allocation to treatments used?
6-was the potential degree of blindness used?
7-was there a satisfactory statement of criteria for outcome measures? Was a primary outcome measure Identified?
8-were the outcome measure appropriate?
9-was a pre-study calculation of required sample size reported?
10-was the duration of post-treatment followed up?
11-were the treatment and control groups comparable in relevant measures?
12-were a high proportion of the subjects followed up?
13-were the dropouts described by treatment and control groups?
14-were the side effects of treatment reported?
15-how were the ethical issues dealt with?
16-was there a statement adequately describing or referencing all statistical procedures used?
17-what tests were used to compare the outcome in test and control patients?
18-were 95% confidence intervals given for the main results?
19-were any additional analyses done to see whether baseline characteristics (prognostic factors) influenced the outcome observed?
20-were the conclusion drawn from the statistical analyses justified?

References

1. Rosen S. Evidence-based medicine is gaining ground with clinicians.BBI newsletter, 2002.
2. Sackett David, Etraus ES, et al. How to practice and teach EBM. Evidence based medicine.2th ed: Churchill Livingstone, 2001.
3- Scketi D, Straus ES. Richardson WS. How to find current best evidence. Evidence based medicine. 2Th ed: Churchill-Livingston, 2001.
4- .Silvermam WA, chalmers, et.al. An appreciation, controlled clinical trials. 1992; 13: 100-5.
5-Jadad AR, Rennie D. The randomized controlled trials gets a Middle-aged checkup. JAMA 1995;279:319-20.
6-Jadad RA. Randomized controlled trials: the basic.BMJ Books,1998.
7-Jadad RA. Randomized controlled trials, A users guide. BMJ Books,1998.
8-Jadad RA.Bias in RCTS: Beyond the sequence generation. BMJ Books,1998.
9-colditz GA. Meller JN, Mosteller F.How study design affect outcomes in comparisons of therapy .Therapy Stat Med 1989;5:441-54.
10-Schulzk F, Grimes DA, et al. Blinding and exclusions after allocation in RCTs.BMJ 1996;312:742-4.
11-Jadad AR, Moore RA, et al.Assessing the quality of reports on RCTs , is blinding nessassry?. Controlled Clin Trails 1996;17:1-12.
12-Moher D, Jadad AR, et al.Assessing the quality of RCTs :An ahhotated biblography of scale cheklists. Controlled Clin Trails 1995;16:62-73.
13-Ohlsson A, Lacy JB, et al. Quality assessment of RCTs:Evaluation by the chalmers vesus Jadad method. Annual Cochrane Calloquium 1995:V9-V10.
14-Bender JS, Halpern SH, et al. Quality and retrieval of obestetrical anesthesia RCTs. Can J Anesth 1997:14-18.