

Evaluating the Internal Validity of a Randomized Controlled Trial

by Michael Turlik, DPM¹ 

The Foot and Ankle Online Journal 2 (3): 5

This paper discusses the important elements to look for when evaluating a randomized controlled trial for internal validity. At the end of the paper the article analyzes a randomized controlled trial found in the first article of this series. This is the second in a series of articles introducing practicing podiatric physicians to evidence-based medicine.

Key words: EBM, evidence-based medicine, randomized controlled trial

Accepted: February, 2009

Published: March, 2009

This is an Open Access article distributed under the terms of the Creative Commons Attribution License. It permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©The Foot and Ankle Online Journal (www.faoj.org)

Is this treatment effective? The best study to answer this question is a randomized controlled trial (RCT). Once a RCT is found to answer the therapeutic question of interest the reader must decide whether or not the authors took all of the important steps to minimize bias. The strength of the inference that we can take from the trial is based upon how well the authors have planned, executed and reported safeguards to minimize bias in their study. The authors could reach an invalid conclusion to the question either because of bias or random error. The evaluation of the safeguards implemented by the authors is referred to as the internal validity of a RCT. Methodological quality is a continuum not a dichotomy and is often in the eye of the beholder.

Even if the authors have described in sufficient detail all of the methods which we would commonly associate with minimizing bias the results of the study may be affected by random error (chance), which is why we can never be sure exactly where the truth may lie. This is the second in a series of articles about evidence-based medicine (EBM) for practicing podiatrists. This article will help podiatric physician in evaluating the internal validity of a RCT.

Bias

Returning to the initial article of this series,¹ you will recall that we had found two articles^{2,3} to help us in evaluating whether the use of insoles with magnets were an effective treatment for reducing pain from diabetic neuropathy. In the initial article of this series a reference was made to a PhD who advocated the use of magnetic insoles for the treatment of symptomatic diabetic neuropathy. In addition, she had developed and was marketing a new type of magnetic insole for this purpose.

Address correspondence to: Michael Turlik, DPM
Email: mat@evidencebasedpodiatricmedicine.com

¹ Private practice, Macedonia, Ohio.

The person presenting this information could be said to be biased toward the use of magnetic insoles specifically her design in the treatment of symptomatic diabetic neuropathy. Bias can be defined as an opinion or feeling that strongly favors one side in an argument or one item in a group or series; predisposition, prejudice.

However, when we consider bias in a RCT, we specifically refer to non-random systematic errors in the design or conduct of the study. Bias is described as any process at any stage of the study which produces results which deviate systematically from the truth. Bias in RCT's is usually not intentional but it is pervasive and insidious. There are many specific types of bias associated with a RCT to include: selection, measurement, and analysis bias. The result of bias in a clinical study is to overestimate the treatment effect as a result, the intervention may appear to work when it really doesn't. Methods in clinical trials which are used to minimize bias include the following: randomization, concealment allocation, blinding, and intention to treat analysis.

Randomization

Case reports and case series while popular in the podiatric literature cannot be used to demonstrate treatment efficacy/effectiveness. These study designs lack a control group and as a result are best used to generate rather than test hypotheses of clinical efficacy/effectiveness. Case-control and cohort studies are observational studies which utilize a control group however, the question always is: Are the two groups similar enough? Is the treatment effect seen in these studies due to the intervention or the difference in prognostic factors between the groups? It is well accepted that non randomized controlled trials exhibit a greater treatment effect than randomized controlled trials.⁴ The only method currently available to provide two groups which are similar for known and unknown prognostic factors is randomization.

Selection bias is the term used to explain the preferential allocation of participants with similar prognostic factors to the same arm of the clinical trial. Random allocation is a study method which minimizes selection bias.

Random allocation can be defined as a process by which all participants have the same chance of being assigned to either treatment arm. One of the first things to evaluate when reading the methods section of a RCT is to determine if the authors of the study generated a randomized sequence consistent with the preceding definition. The following methods are *not* consistent with random allocation methods: date of birth, hospital chart number, alternate selection, date of entry, or days of the week. Unlikely methods of random allocation in clinical trials are: flipping a coin, rolling dice, or choosing colored balls out of a bag. Typically random allocation in clinical trials will be achieved by the use of a random number table or a computer-generated series of numbers. Another question which should be addressed in the methods section by the authors is *who* generated the randomization sequence? It is best when an independent third-party not associated with the study generates the sequence. In a study of RCT's published in popular podiatric medical journals only two of the nine studies described a random allocation process.⁵

Concealment allocation

After the randomization sequence is generated the list may be given to the investigator responsible for enrolling participants in the study. This is referred to as unconcealed participant allocation. The investigator may steer participants to certain treatment arms based upon prognostic factors either consciously or unconsciously. Concealment allocation can be defined as the process by which the physician is blinded to the randomized sequence which was generated. The person who enrolls participants in the trial should not be the same person who generates the allocation sequence.

In RCTs where concealment allocation has not been utilized there is an over estimation of treatment effect compared to trials which conceal the allocation sequence. The increase treatment effect may be 20 to 30%.⁶ The average bias associated with lack of adequate concealment allocation was less for outcomes which were objectively evaluated (death, ulcer closure) than subjectively evaluated (pain, patient reported outcomes).

The description the authors used for concealment allocation is usually found in the methods section of the paper. For example, a common description might be "...a neutral third party has generated a series of sequentially numbered opaque sealed envelopes (SNOSE) containing the randomization sequence for each participant to be opened at the time the investigator enrolls the participant in the study." As a result the investigator is blinded to the treatment arm to which the participant will be enrolled.

It is very common today to have a centralized allocation process when the investigator enrolls a participant in the study he or she will call an off-site location to determine the next group to which the subject is enrolled. A similar system is also used by computer over the internet. Sometimes the allocation sequence may be kept in the pharmacy which will be contacted by the investigator prior to enrolling the participant. Adequate concealment allocation helps to limit selection bias. In a study of RCTs published in common podiatric medical journals none of the nine studies described a process for concealment allocation.⁵

Were patients in the treatment and control group similar with respect to known prognostic factors? Problems with the randomization allocation sequence and the concealment allocation process may result in an imbalance in baseline prognostic factors. If both of these steps have been followed and there still remains an imbalance in some important prognostic factor it will be assumed that this is due to chance rather than bias. The larger the study the less likely this is to occur.

Blinding

Blinding in a clinical trial can be defined as withholding information about treatment allocation from those who could potentially be influenced by this information. Un-blinded studies exhibit an increased treatment effect compared blinded studies.⁷ In the methods section the authors should describe in some detail who was blinded, how they were blinded, and the success of blinding.

Who was blinded? Certainly participants and investigators can be blinded. Less commonly recognized is that data collectors and analysts should be blinded. Participants should be blinded because they may use other effective interventions, may report symptoms differently, or may drop out if they perceived they have received a placebo therapy. Investigators should be blinded because they may prescribe effective co-interventions, influence follow-up, or patient reporting. Data collectors and analysts should be blinded because they may exhibit differential encouragement during performance testing, exhibit variable recordings of outcomes, or differential timing and frequency of outcome measurements.

How was blinding achieved? In the case of medication the placebo should be the same size, shape, color and taste as the therapeutic intervention. When using a gold standard a double dummy process should be employed. When using a sham procedure both instruments should look alike, sound alike, have the same lights and duration. Separate waiting rooms may be necessary for each treatment arm to prevent interactions between groups. Sometimes the therapeutic intervention under investigation precludes the investigators and/or participants from being blinded however, it is difficult to understand why data collectors and analysts cannot be blinded.

There is no universal agreement upon how to assess blinding⁸ or even if it should be assessed. Often times study authors will ask investigators and participants to guess at their treatment allocation and report the results. Some would suggest looking for bias generating consequences instead of contamination and co-interventions.

Measurement bias is defined as inaccurate measurement due to either the accuracy in the measurement instrument or bias based upon study expectations of participants and investigators. Blinding will help to limit measurement bias. In a study of RCT's published in popular podiatric journals only two of the nine trials described a process for blinding.⁵

Intention to Treat Analysis

Intention to treat (ITT) analysis can be defined as the strategy for the analysis of randomized controlled trials to ensure all participants are compared in the groups to which they were originally randomly assigned. Although this sounds simple it is difficult to understand and often confused in the literature. In general, all patients who were randomized at the beginning of the trial must be accounted for during the analysis. Certainly there are some exceptions⁹ however, failure to account for all the participants at the conclusion of the trial will result in analysis bias, overestimating the treatment effect. ITT preserves the prognostic balance in the treatment arms achieved by randomization and increases generalizability.

During the course of the study participants may elect not to participate after they were randomized or change treatment arms for various reasons. In ITT analysis they should always be analyzed into the group they were originally allocated regardless of the treatment received. In addition, during the course of the study participants may be lost to follow-up.

It is well-established that participants who drop out of a study have a different prognosis than those who remain.¹⁰ There are many different methods to infer or impute lost study results for example, last value carried forward, or worse case scenario. Unfortunately these are only estimates for missing data. ITT prevents conscious or unconscious attempts to influence study results by excluding participants.

Results in a RCT may be presented in two different manners. Per protocol indicating that the results only include patients who have successfully completed the trial. ITT analysis will be reported indicating that all participants are accounted for. Per protocol analysis answers the question what will happen if my patients all comply with the treatment intervention (explanatory). Intention to treat analysis answers the question what will happen in real life using this treatment intervention (pragmatic). If there is a difference between per protocol analysis and ITT analysis, the loss of follow-up has been large and the inference from the study is reduced. Intention to treat analysis is a more pragmatic and a more conservative estimate of treatment effect and minimizes a type I reporting error.

When reviewing the methods section of a RCT to evaluate intention to treat analysis the reader should look for how intention to treat was performed. It is unacceptable for the authors simply to state that the data was analyzed on an intention to treat principle without explanation. In a study of RCT's published in common podiatric journals four of the nine papers reported that the data was analyzed on an intention to treat basis.⁵

Magnets in the treatment of diabetic neuropathy

In the first article of this series¹, a RCT⁵ was identified which evaluated the usefulness of magnetic insoles in reducing the symptoms of painful diabetic neuropathy.

Using the information presented earlier the internal validity of this article was critically analyzed.

Randomization: The authors in the methods section describe an equal random allocation procedure utilizing a computer. It is unclear who actually generated the sequence and how it was accomplished. In addition, randomization was stratified by center and gender.

Concealment allocation: In the methods section the authors report that neither the participants or investigators were aware of the treatment allocation. However, they did not elaborate as to the method. It may be that utilizing the computer a centralized allocation process was used.

Baseline comparison: Although the active treatment group and the sham group appear to be similar with regards to baseline characteristics, the groups appear to be dissimilar when baseline outcome measurements are compared. The baseline outcome measures for the sham group appear to be worse than the intervention group.

Blinding: The authors state that the sham and active magnetic insoles were identical with regards to the appearance, consistency and weight. In the event the insole did not fit the shoe and needed to be trimmed the authors described a process by which an uninvolved third party would make the adjustments. In addition, all data was submitted blindly to an uninvolved third-party for analysis. In the results section the authors report their efforts on assessing the effectiveness of the methods used in blinding the investigators and participants. There was no indication that any contamination or co-interventions were detected. However, in the methods section there was no indication that the authors were attempting to measure contamination or co-interventions.

Intention to treat analysis: The results of the study were not analyzed on an ITT basis. Furthermore, the author states that participants with incomplete data were excluded from analysis.

In the results section the authors discussed dropouts and their decision not to analyze the data on an intention to treat basis. The authors chose to use four different primary outcomes each with different numbers of participants.

Summary of internal validity: Based upon the methods and results section of the paper it is clear that the authors attempted and succeeded in blinding participants, investigators and data analysts. It is less clear as to the method of random allocation and concealment allocation. The authors made no attempt to analyze the data utilizing the intention to treat principle.

References

1. Turlik M: Introduction To Evidence-based Medicine. *Foot and Ankle Journal* 2: 4, 2009.
2. Pittler M, Brown EM, Ernst E: Static magnets for reducing pain: systematic review and meta-analysis of randomized trials. *CMAJ* 177: 736 – 742, 2007.
3. Weintraub MI, Michael I. Weintraub, Wolfe GI, Barohn RA, Cole SP, Parry GJ, Hayat G, Cohen JA, Page JC, Bromberg MB, Schwartz SL, Magnetic Research Group: Static magnetic field therapy for symptomatic diabetic neuropathy: a randomised, double-blind, placebo-controlled trial. *Arch Phys Med Rehabil* 84 (5): 736 – 746, 2003.
4. Moore A, McQuay H: *Bandolier's Little Book of Making Sense of the Medical Evidence*. Oxford University press, Oxford England, 2006.
5. Turlik M., Kushner D, Stock D: Assessing the Validity of Published Randomized Controlled Trials in Podiatric Medical Journals. *JAPMA* 93 (5): 392 – 398, 2003.
6. Wood L: Empirical evidence of bias in treatment effect estimates in controlled trials with different meta-epidemiological study interventions and outcomes *BMJ* 336 : 601 – 605, 2008.
7. Poolman R: Reporting of Outcomes in Orthopaedic Randomized Trials: Does Blinding of Outcome Assessors Matter? *J Bone Joint Surg* 89A: 550 – 558, 2007.
8. Hróbjartsson A. trials taken to the test: an analysis of randomized clinical trials that report tests for the success of blinding. *International Journal of Epidemiology* 36 (3): 654 – 663, 2007.
9. Fergusson D, Aaron SD, Guyatt G, Hébert P: Post-randomisation exclusions: the intention to treat principle and excluding patients from analysis. *BMJ* 325: 652 – 654, 2002.
10. Altman D: Clinical trials. In: *Practical statistics for medical research*. London: Chapman & Hall, 1991.