

How to Interpret External Validity of a Randomized Controlled Trial

by Michael Turlik, DPM¹ 

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

The fourth and final article discussing the interpretation of randomized controlled trials for podiatric physicians involves the critical analysis of external validity. External validity is defined, examples from the foot and ankle literature are reviewed and a critical analysis is explained for the randomized controlled trial involving the use of magnetic insoles in the treatment of symptomatic diabetic neuropathy. This is part of an ongoing series of articles about evidence-based medicine to assist podiatric physicians develop the knowledge, skills and values to be successful in a changing practice environment.

Key words: Evidence-based medicine, external validity, randomized controlled trial

Accepted: April, 2009

Published: May, 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)

When evaluating a randomized controlled trial (RCT) the first consideration is internal validity. Did the authors take appropriate measures to minimize bias? If the reader is satisfied that the authors did plan, implement and report methods which would minimize bias the reader would then evaluate the results of the trial. If the trial results revealed a clinically important difference between treatment arms the final question the reader will have to answer is; can I and should I apply the results of this study to my practice? This is a matter of judgment rather than statistical analysis. It is often referred to as relevance or external validity of the trial. This is the fourth and final article in this series of articles for podiatric physicians designed to help them understand, interpret and implement RCTs in their practices.

Types of Randomized Controlled Trials

Explanatory trials are characterized by having strict inclusion criteria, highly homogenous study groups, usually only present the results as per protocol and often use placebos as the comparator. Explanatory trials answer the question does this intervention work, this relates to the efficacy of the intervention. **Pragmatic** trials have increased variability of the participants, use an active comparator (gold standard) and evaluate the results as per protocol and intention to treat analysis. Pragmatic trials evaluate the effectiveness of an intervention; does this treatment work in real life? These are two ends of a continuum, the more pragmatic the trial is the better its external validity.

External validity or relevance involves the interpretation of the trial relative to the **readers** practice environment. Issues to be resolved by the podiatric physician when evaluating a trial for external validity are: trial participants, location of the study, intervention, outcomes and harms described. In general, RCTs are better suited for answering the question; Will this treatment work, rather than is it worth it?

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

¹ Private practice, Macedonia, Ohio.

Trial Participants

The first step for evaluating the trial participants is to review the inclusionary and exclusionary criteria listed by the author in the methods section of the article. For example, in a recent study¹ the authors looked at the use of duct tape in the treatment of pedal verruca in children 4-12 years of age. Can the results of this study be generalized to adults? In another recent study² of **midportion** Achilles tendon pain, all participants were confirmed by ultrasound prior to inclusion. Should the results be generalized to patients with **insertional** Achilles symptoms and should all patients be confirmed by ultrasound prior to beginning treatment as described in the study?

Are the participants in the trial similar with respect to age, gender, co morbidities and severity of disorder? It is easy to rationalize that the patient(s) in the podiatric physicians practice differs from those in the trial regarding **some** demographic aspect. "My patients are slightly older/younger, have more more/less co morbidities, take multiple/ less medications, speak/don't speak English, wear the same /different shoes, have the same/different occupation." Would my patient have been included in the trial? A good rule of thumb is that if the answer to this question is yes the results probably are relevant to your patient (s).

Location of the Study

In this classic study³ the authors evaluated a clinical maneuver to detect infected bone in hospitalized patients with diabetic foot ulcers. Can the results be generalized to an office setting where the ulcers may be less severe than the participants in the study? Another study⁴ looked at pulse versus continuous terbinafine for onychomycosis. The setting was a VA (What does this mean?) hospital in Minneapolis. Should the results be generalized to an upper income suburban private practice?

A common concern regarding study location is; are the results of a trial done in a tertiary care center by specialists relevant to primary care settings?

Another concern would be; are differences in the healthcare delivery system of other countries relevant to podiatric physicians practicing in the United States? The question the podiatric physician needs to answer is; was the setting of this study similar **enough** to my practice so that the results can be generalized?

Intervention

Is the intervention proposed in the study feasible for me? Am I technically proficient to perform the procedure? Is the device available to me? How much does the intervention cost and who will pay for it? All important considerations before deciding if the intervention is relevant to the individual podiatric physicians practice. The results of a randomized controlled trial may demonstrate a clinical benefit however; the intervention may not be relevant to the podiatric physicians practice.

The methods section of the paper should describe in enough detail the intervention and comparator to allow the reader to make an informed judgment. In a study⁵ of foot orthoses to treat plantar fasciitis the reader needs to determine if the method of casting and fabrication of the foot orthoses described are similar enough to his /her practice to allow for generalization of the results. A different study⁶ evaluated the difference between surgery, foot orthoses and watchful waiting for symptomatic hallux valgus. In addition to the question again regarding foot orthoses, the reader has to decide if the procedure performed is similar enough to the typical surgical procedure performed for hallux valgus in his/her practice. In an earlier article of this series⁷ it was shown that Graftskin was a more effective treatment for diabetic foot ulcers when compared with moistened saline gauze dressings.⁸ The estimated cost for a single treatment with Graftskin varies from \$1000 - \$1200.⁹ Despite the high cost of Graftskin a recent economic analysis demonstrated that for venous stasis ulcers Graftskin was more effective and less costly than treatment with unna boots.¹⁰

Outcomes

Is the outcome meaningful to my patients? Patient centered outcomes are the hallmark of evidence-based medicine. Validated patient centered outcomes are the most relevant to clinical practice. The authors should describe the primary outcome(s) in the methods section and reference validation efforts.

In a recent study¹¹ the authors looked at the use of Alendronate in acute Charcot arthropathy. The authors measured changes in the following to evaluate the success of the intervention: serum collagen COOH-terminal telopeptide of type 1 collagen (ICTP), osteocalcin, testosterone, estradiol, thyroid hormones, parathyroid hormone, follicle-stimulating hormone, leutinizing hormone, IGF-1, calcitriol, urinary hydroxyprolin, serum alkaline phosphatase. These are examples of surrogate outcomes. An important question to ask regarding surrogate outcomes is; is there a strong, independent, consistent association between the surrogate outcome and a patient-important outcome?¹²

In a trial comparing two different surgical procedures¹³ for the correction of hallux abducto valgus deformity the authors used multiple outcomes. They consisted of: the American Orthopedic Foot and Ankle Society clinical rating scale for deformities of the hallux, EuroQol, visual analogue scale, and radiographic changes. Some of the outcomes are patient oriented (VAS, EuroQol) some are physician centered outcomes (range of motion, radiographic findings), some are validated others are not validated. Physician centered outcomes, and surrogate outcomes are not as relevant to clinical practice as patient centered outcomes. Composite outcomes may be important however, they need to be cautiously reviewed to ensure the results are not misleading.¹⁴

The question the podiatric physician needs to answer is; is the outcome used in this study patient centered, and validated? If the answer to this question is yes then the podiatric physician should feel comfortable in generalizing the results to his/her patient(s).

Harms

Were the benefits from the intervention in the trial worth the potential harms that occurred? RCTs present a less biased estimate of conceivable harmful effects than other study designs due to balancing of known and unknown prognostic factors. However, randomized controlled trials are unlikely to enroll enough participants to demonstrate rare and serious adverse effects which may be evident in the larger population. In addition, unlike reporting of efficacy in randomized controlled trials the reporting of harms is thought to be inadequate and incomplete.¹⁵ **Larger observational studies are usually better to assess harms from therapeutic interventions than RCT's.** In a study evaluating pharmaceutical versus non pharmaceutical trials in rheumatoid arthritis the authors found that pharmaceutical trials reported data about adverse effects on harms more often than non-pharmaceutical trials.¹⁶ Fewer than half of the non pharmaceutical trials reported any harms.

Although not always reported in the methods section the authors should describe the method by which information on harms were collected and reported. Ideally, in the results section the author should report all of the adverse outcomes which occurred, the frequency with which they occurred and if important the time which they occurred. Consider the number of participants who withdrew from the trial due to adverse effects, and the completeness of the reporting of the adverse effects when evaluating a paper for harms.

In an article evaluating duct tape for pedal verruca¹ in children referenced earlier the authors briefly described in the methods section a passive questionnaire used to elicit information about adverse effects. Adverse effects are described both in the narrative section and a table in the results section. Three participants withdrew from the duct tape arm of the trial and no participants withdrew from the placebo arm of the trial due to adverse effects. Participants using the intervention demonstrated a larger percentage of localized adverse skin effects which were described as mild.

In another article referenced earlier¹³ comparing two different surgical procedures for hallux abducto valgus deformity the author does not discuss the evaluation of harms in the methods section. However in the results the authors discuss complications occurring during the study which are limited to one surgical revision due to recurrence one year post operative. One participant developed an asymptomatic nonunion which resolved during the one-year follow-up period. Apparently none of the 100 participants developed a postoperative infection.

Was there a difference in harms between treatment arms? Was the difference statistically/clinically significant? Were there any serious adverse consequences noted between groups? Although not ideal for investigating harms any information on adverse effects presented in the paper should be used by the podiatric physician in deciding if a clinically significant result should be implemented in his or her practice. The information on harms should also be communicated to patients prior to beginning the intervention.

Use of magnetic insoles for diabetic peripheral neuropathy

The initial article of this series¹⁷ identified a RCTs¹⁸ which would assist us in evaluating the usefulness of magnetic insoles in the treatment of symptomatic diabetic neuropathy. The following two articles^{7,19} critically evaluated the internal validity and results of the identified article. Using the information presented earlier in this article we will critically evaluate the external validity of the RCT evaluating magnetic insoles for diabetic neuropathy.

Trial Participants

The authors clearly explained the inclusionary and exclusionary criteria necessary for the participants to enter the trial. The entry criteria appear to be rather broad increasing the external validity of the trial.

The only concern would be whether or not magnetic insoles would be indicated as a first-line treatment since the participants were required to have symptoms which were constant and present over six months and which were refractory to various medications.¹⁸

Location of the Study

Three hundred and seventy five participants were recruited from 48 sites in 27 states. Participants were recruited from various specialties to include neurology, podiatry, diabetic clinics and other private practices. This multicenter study appears to have high external validity with regards to the location of the study.

Intervention

The authors provided a thorough technical description of the magnetic insole used in the study. The magnetic insoles used in the study are readily available over the Internet for less than \$100 per pair. Whether and how these magnetic insoles differ from other commercially available magnetic insoles was not addressed in the study.

Outcomes

In the methods section the authors describe the use of two different instruments to measure four different primary outcomes. An 11 point visual analogue pain scale (0-10) was used to measure numbness and tingling/burning pain. A quality of life (QOL) instrument was used to measure foot pain and sleep interruption. Visual analogue scales and quality of life instruments are considered patient centered. In the methods section no references were cited regarding validation of the instruments furthermore, the quality of life instrument was not identified.

Harms

Under the methods section the authors stated that adverse effects would be monitored and reported. The adverse effects the authors were looking for included ulceration, abrasion, allergic reaction or infection. The authors reported in the results section no complications from either trial arm. However, in the flowchart of the trial reference was made to six participants in the intervention arm and four participants in the placebo arm experiencing complications.

Summary

The authors described in sufficient detail methods which would minimize bias with the exception of intention to treat analysis. The results of the study were not clearly presented comparing magnetic insoles and placebo. Using effect sizes calculated from the original data a small clinically significant effect was obtained with one of the four primary outcomes measured. The external validity of the study seems high with regards to trial participants, location of the study and the intervention. The reader will need to decide whether the cost of the intervention is worth the questionable efficacy of the intervention.

References

1. de Haen M, Spigt MG, van Uden CJT, van Neer P, Feron FJM, Knottnerus A: Efficacy of duct tape vs placebo in the treatment of verruca vulgaris (Warts) in primary school children. *Arch Pediatr Adolesc Med* 160: 1121 – 1125, 2006.
2. Petersen W, Welp R, Rosenbaum D: Chronic Achilles tendinopathy: A prospective randomized study comparing the therapeutic effect of eccentric training, the AirHeel brace, and a combination of both. *Am J Sports Med* 35: 1659 – 1667, 2007.
3. Grayson ML, Gibbons GW, Baloh K, Levin E, Karchmer AW: Probing to bone in infected pedal ulcers: a clinical sign of underlying osteomyelitis in diabetic patients. *JAMA* 273 (9): 721 – 723, 1995.
4. Warshaw E, Fett D, Bloomfield H, Grill J, Nelson D, Quintero V, Carver S, Zielke G, Lederle F: Pulse versus continuous terbinafine for onychomycosis: A randomized, double-blind, controlled trial *Journal of the American Academy of Dermatology*. 53 (4): 578 – 584, 2002.
5. Landorf K, Keenan A-M, Rushworth RL: Effectiveness of foot orthoses to treat plantar fasciitis. *Arch Intern Med* 166: 1305 – 1310, 2006.
6. Torkki M, Malmivaara A, Seitsalo S, Hoikka V, Laippala P, Paavolainen P: Surgery vs. orthosis vs watchful waiting for hallux valgus: A randomized controlled trial. *JAMA* 285:2474 – 2480, 2001.
7. Turlik M. How to interpret the results of a randomized controlled trial. *Foot and Ankle Online Journal*. 2: 4, 2009.
8. Veves A, Falanga V, Armstrong DG, Sabolinski ML, Apligraf Diabetic Foot Ulcer Study: Graftskin, a human skin equivalent, is effective in the management of noninfected neuropathic diabetic foot ulcers. *Diabetes Care* 24: 290 – 295, 2001.
9. Hanft J, Williams A, Kyramarios C, Temar K: Are tissue replacements cost effective? *Podiatry Today* 16: 2003.
10. Schonfeld WH, Villa KF, Fastenau JM, Mazonson PD, Falanga V: An economic assessment of Apligraf® (Graftskin) for the treatment of hard-to-heal venous leg ulcers *Wound Repair and Regeneration* 8: 251 – 257, 2001.
11. Pitocco D, Pitocco, Ruotolo V, Caputo S, Mancini L, Collina CM, Manto A, Caradonna P, Ghirlanda G: Six-month treatment with alendronate in acute charcot neuroarthropathy. *Diabetes Care* 28: 1214 – 1215, 2005.
12. Guyatt G, Drummond R, Meade M, Cook D. *Users' guides to the medical literature*. New York: McGraw-Hill Medical, 367 – 374, 2008.
13. Saro C, Andrén B, Wildemyr Z, Felländer-Tsai L: Outcome after distal metatarsal osteotomy for hallux valgus: A prospective randomized controlled trial of two methods. *Foot and Ankle International* 28: 778 – 787, 2007.
14. Montoti V, Montori VM, Busses JW, Miralda GP, Ferreira I, Guyatt GH: How should clinicians interpret results reflecting the effect of an intervention on composite end points: should I dump this lump? *EBM* 10: 162, 2005

15. Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, CONSORT Group: Better reporting of harms in randomized trials: An extension of the CONSORT statement. *Annals of internal medicine* 141: 781 – 788, 2004.
16. Ethgen M, Boutron I, Baron G, Ethgen M: Reporting of harm in randomized controlled trials of nonpharmacologic treatment for rheumatic disease. *Ann Intern Med* 143: 20 – 25, 2005.
17. Turlik, M. Introduction To Evidence-based Medicine. *Foot and Ankle Online Journal*. 2:4,2009
18. Weintraub MI, 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: 736 – 746, 2003.
19. Turlik M. Evaluating the internal validity of a randomized controlled trial. *Foot and Ankle Online Journal*. 2: 5, 2009.